



D2I DATA TO INFORMATION

Higher order of perception and analysis to recognize regularities

A GLIAL APPROACH

➤ Significant Development Example (10/03/04)

Merck & Co., Inc. announced a voluntary worldwide **withdrawal** of **VIOXX** (rofecoxib), its arthritis and acute pain medication. Merck's decision, which is effective **immediately**, is based on new, three-year data from a prospective, randomized, placebo-controlled clinical **trial**. _____

➤ Significant Development Example (10/11/04)

Merck & Co., Inc. begins looking outside for possible **successor** to **CEO**. The Wall Street Journal reported that the Merck & Co., Inc. Board has begun **looking outside** for a possible successor to Raymond Gilmartin, the Company's longtime Chairman and Chief Executive Officer. The move comes amid a string of **bad news** at the drug maker. _____

What's more telling — **highlighted** words or others? more useful for analysis?

Are these complex?

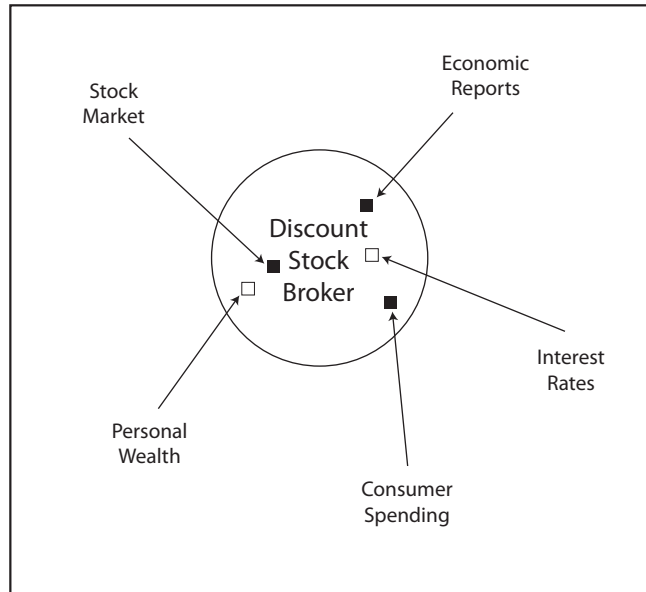
If we can find no simple features, we tend to lose interest — simple, easy to describe features are aesthetically pleasing...

Are these random?

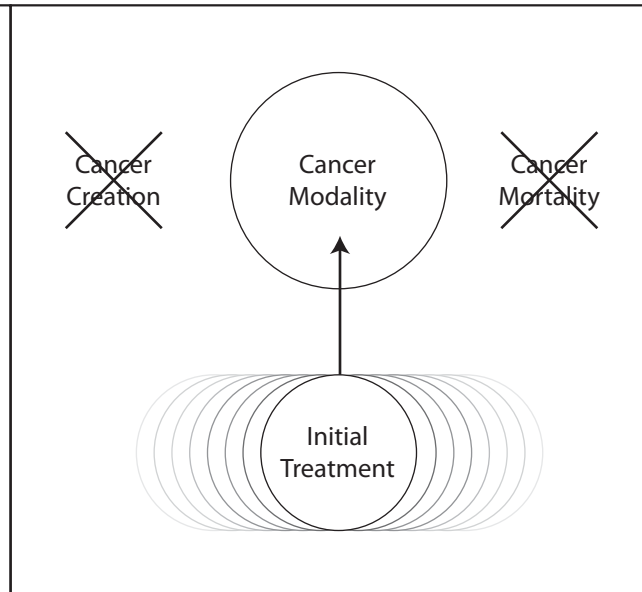
When we find irregularities, it implies redundancy is not present, which means a shorter description cannot be given....

RANDOMIZATIONS

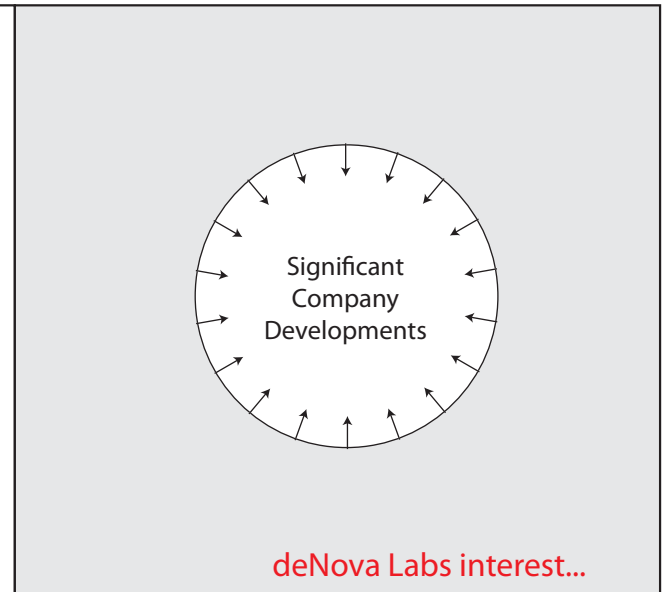
Randomness from environment



Randomness from initial conditions



Randomness from intrinsic generation



- random input from a random outside
- every square controlled by outside
- non-repeatable randomness

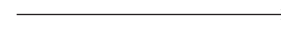
- random input from a fixed outside
- random input only at the beginning
- non-repeatable randomness

- no random output from outside
- start with the same black square
- repeatable randomness

customer potential value?



patient biographic profiling



Stochastic/Monte Carlo/Random Walk

Chaos Theory

Cellular Automata/Autoplectic

CONSTRUCTING MODELS

- If the behavior of a system is complex, does the model for the system have to be complex?
No...
 - Think in terms of programs and not traditional equations
 - Identify which effects are important enough that they must be kept
 - Use experiments to essentially isolate each individual effect
 - Explicitly measure each criteria's importance
- Cannot extract a small set of numbers from the observed behavior of the system, and then see how accurately these numbers can be reproduced by a model
- Must compare overall pictures of a system with pictures of a model
- Model cannot be complex or patched to account for new observations (no bolt-ons)
- Get an abstract representation of effects that are important in determining the behavior of a system — model should not operate like system itself

SEQUENCING

Merck & Co., Inc. announced a voluntary worldwide withdrawal of VIOXX (rofecoxib), its arthritis and acute pain medication. Merck's decision, which is effective immediately, is based on new, three-year data from a prospective, randomized, placebo-controlled clinical trial.

MARKER		→ BLOOM					PAST TENSE?					ALL CAPS?					
0	1	2	2	3	1	9 _x	1	9	9	10	2	5 _x					
										← BLOOM		MARKER		→ BLOOM			
9	1	3	9	3	5	4	10	1	0	8	1	5	2				
PUNCTUATION?																	
9		11	1	2	5	2	3	1	10 _x	4	4	1	11			1	
HYPHENATED WORDS?																	
10	1		18 _x			8	5	1									

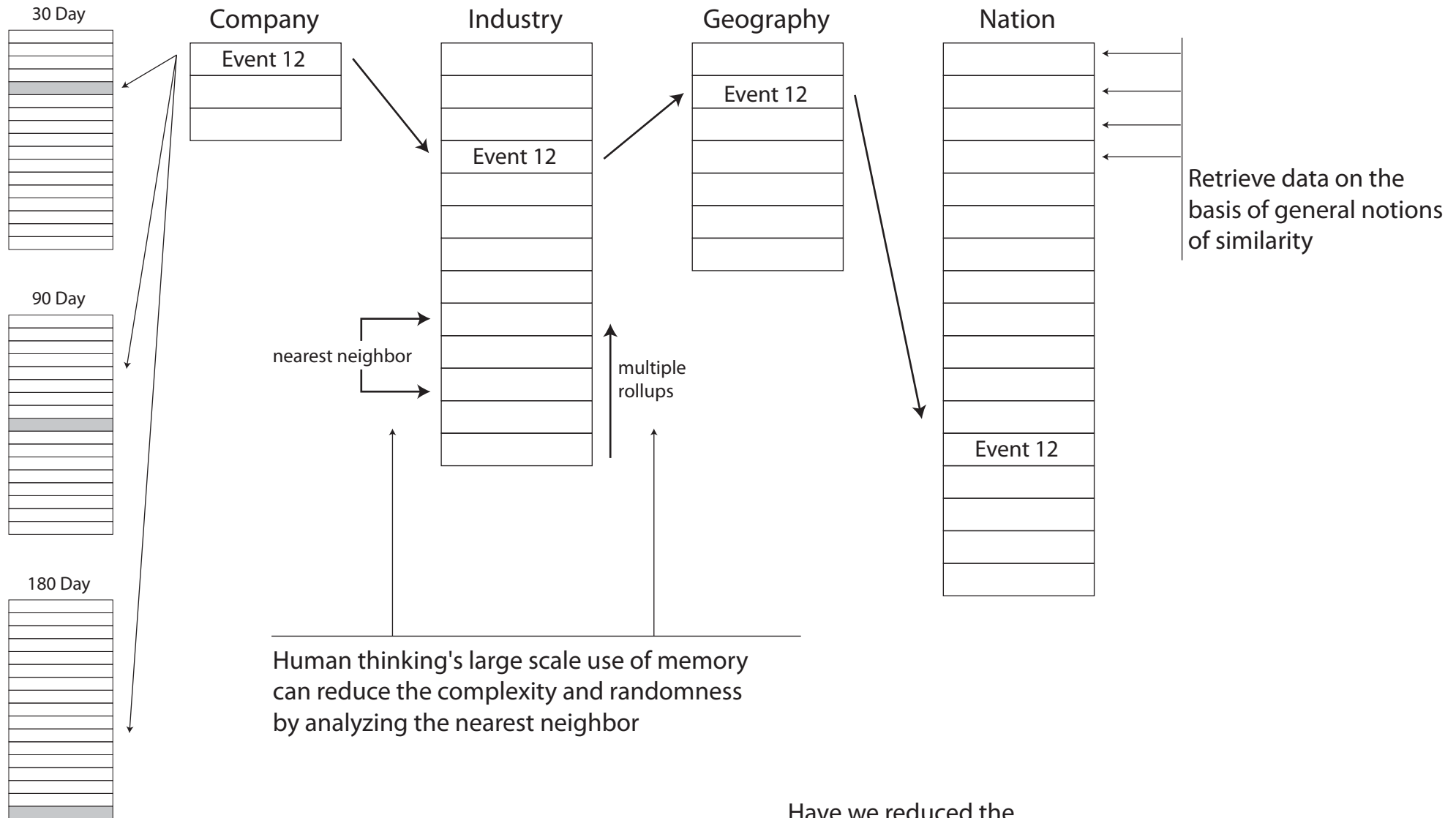
Are these attributes important?

- ¶ starts with sanitized company name? second sentence also?
- Complexity of word after sanitized name?
- Presence and frequency of past, present and future tenses?
- Location of sanitized name relative to punctuation?
- Length of the ¶? Rarity of the subject matter (taxonomy)?

Is this data irregular? Can simple rules find regularities?

Key is not to think about these numbers in size, but in terms of digit sequencing...

VISIBLE PERCEPTIONS



Have we reduced the randomness and complexity?

