I Throw Itching Powder at Tulips

Richard P. Gabriel

IBM Research

Only mystery enables us to live

–Federico Garcia Lorca

Only mystery enables us to live

Only mystery

–Federico Garcia Lorca

So...

- I have rarely written software that was used
- commercially, and when I did, only the loosest
- notion of specification was in place.
- Anyhow, that period extended about 3 months,
- so it's a vanishingly small part of my career.

- What I do is write software as part of doing
- science. I use software as a machine or
- instrument to explore how the mind / brain
- might work.

Software Engineering

- is for producing something you can describe,
- either using specifications (the old-fashioned way)
- or a product backlog (or some other suchlike
- thing in an agile setting). I don’t do anything
- like that.

Let Me Explain the Project I Am Working On

- Social Media in Strategic Communication (SMISC)
- DARPA-BAA-11-64

- Goal 1: Detect, classify, measure, and track
  - the formation, development, and spread of ideas & concepts (memes)
  - purposeful or deceptive messaging and misinformation.

- Goal 2: Recognize persuasion campaign structures and influence operations across social media sites and communities.

- Goal 3: Identify participants and intent, and measure effects of persuasion campaigns.

- Goal 4: Counter messaging of detected adversary influence operations.

- I work on Goal 4. In particular I work on understanding the “adversary,” their roles, their interests and intentions, and I am supposed to formulate a rhetorical strategy and then execute it in natural language.
- Where I mean I am working on making a computer do these things.

DARPA's Example

For example, in one case rumors about the location of a certain individual began to spread in social media space and calls for storming the rumored location reached a fever pitch. By chance, responsible authorities were monitoring the social media, detected the crisis building, sent out effective messaging to dispel the rumors, and averted a physical attack on the rumored location. This was one of the first incidents where a crisis was (1) formed, (2) observed and understood in a timely fashion, and (3) diffused by timely action, entirely within the social media space.

–DARPA-BAA-11-64

Friday, October 18, 13

Programmers mediate between
the negotiated and uncertain truths of business
and the crisp, uncompromising domain of bits
and bytes and higher constructed types.

–Kevlin Henney, *97 Things Every Programmer Should Know*, 2010

We Learn To Do What We Are Told in Programming

Develop a function that when given an initial amount of money (called the principal), a simple annual interest rate, and a number of months will compute the balance at the end of that time. Assume that no additional deposits or withdrawals are made and and that a month is 1/12 of a year. Total interest is the product of the principal, the annual interest rate expressed as a decimal, and the number of years.

–Felleisen et al, *How To Design Programs*

It Is Reinforced

I think the biggest mistake we make with the starting point for undergraduate education is that we introduce programming at all. The right starting point, IMHO, is requirements and specification together with the associated mathematics that they require.

–anonymous, *ruminating on a first course*

Vision and Scope Document

1. Business Requirements

1.1 Background, Business Opportunity, and Customer Needs

A majority of Process Impact employees presently spend an average of 60 minutes per day going to the cafeteria to select, purchase, and eat lunch. About 20 minutes of this time is spent walking to and from the cafeteria, selecting their meals, and paying for their meals by cash or credit card. When employees go out for lunch, they spend an average of 90 minutes off-site. Some employees phone the cafeteria in advance to order a meal to be ready for them to pick up. Employees don't always get the selections they want because the cafeteria runs out of certain items. The cafeteria wastes a significant quantity of food that is not purchased and must be thrown away. These same issues apply to breakfast and supper, although far fewer employees use the cafeteria for those meals than for lunch.

Many employees have requested a system that would permit a cafeteria user to order meals on-line, to be delivered to a designated company location at a specified time and date. Such a system would save those employees who use the service considerable time and it would increase the chance of them getting the food items they prefer. This would improve both their quality of work life and their productivity. Knowing what food items customers want in advance would reduce wastage in the cafeteria and would improve the efficiency of cafeteria staff. The future ability for employees to order meals for delivery from local restaurants would make a wider range of choices available to employees and provide the possibility of cost savings through volume purchase agreements with the restaurants. It might also permit Process Impact to have the cafeteria handle only individual lunches, relying on restaurants to fill orders for breakfasts, dinners, special events, and weekend meals.

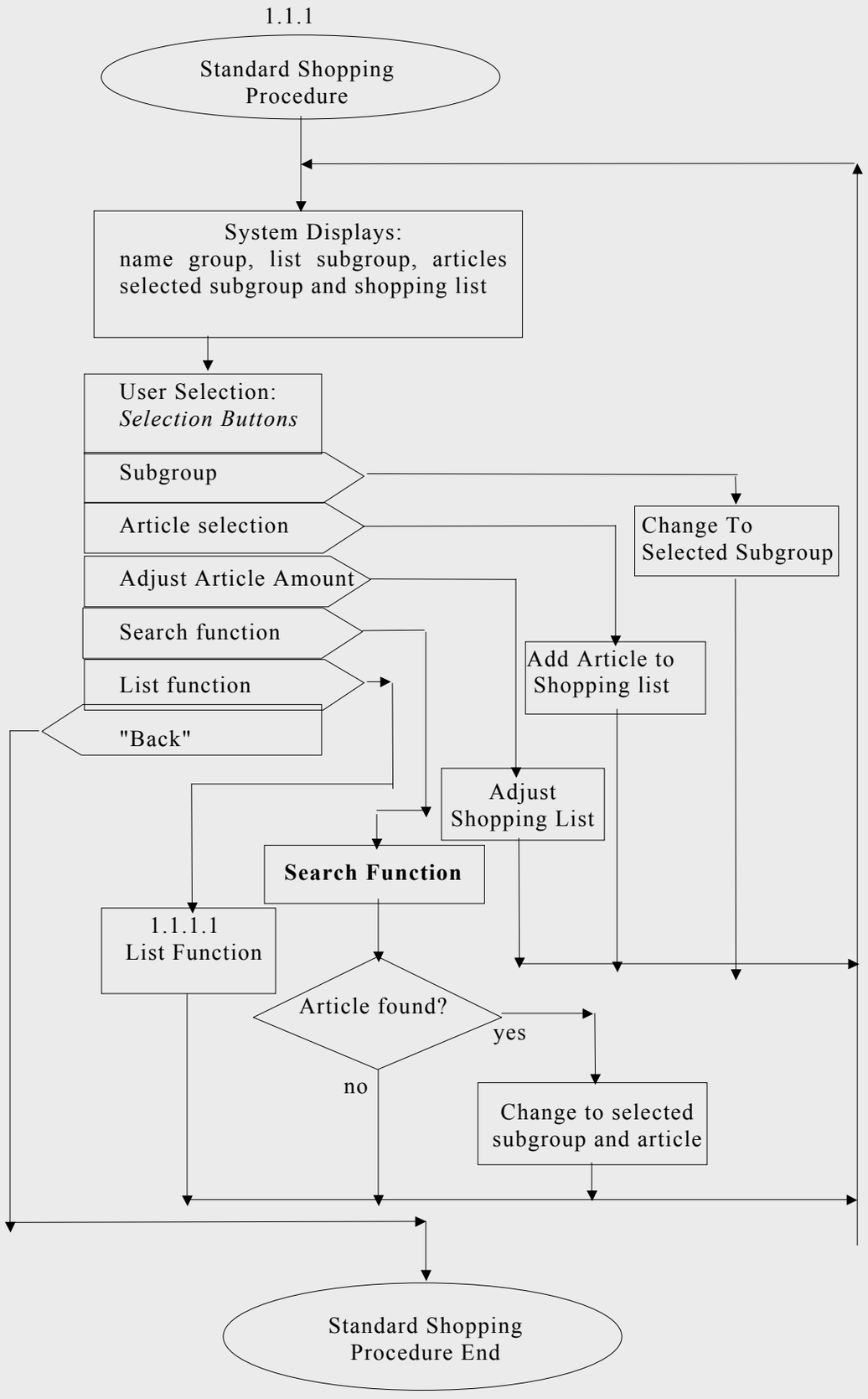
1.2 Business Objectives and Success Criteria

BO-1: Reduce cafeteria food wastage by 50% within 6 months following initial release.[1]
Scale: Value of food thrown away each week by cafeteria staff
Meter: Examination of Cafeteria Inventory System logs
Past [2002, initial study]: 30%
Plan: Less than 15%
Must: Less than 20%

BO-2: Reduce cafeteria operating costs by 15% within 12 months following initial release.

1. This example shows the use of Planguage as a way to precisely state a business objective or other requirement.

Back When “Programming” Was “Coding”


1.1.1
Standard Shopping Procedure
System Displays:
name group, list subgroup, articles
selected subgroup and shopping list
User Selection:
Selection Buttons
Subgroup
Article selection
Adjust Article Amount
Search function
List function
"Back"
Change To Selected Subgroup
Add Article to Shopping list
Adjust Shopping List
Search Function
1.1.1.1 List Function
Article found?
yes
no
Change to selected subgroup and article
Standard Shopping Procedure End

You Think Scrum is Any Better?

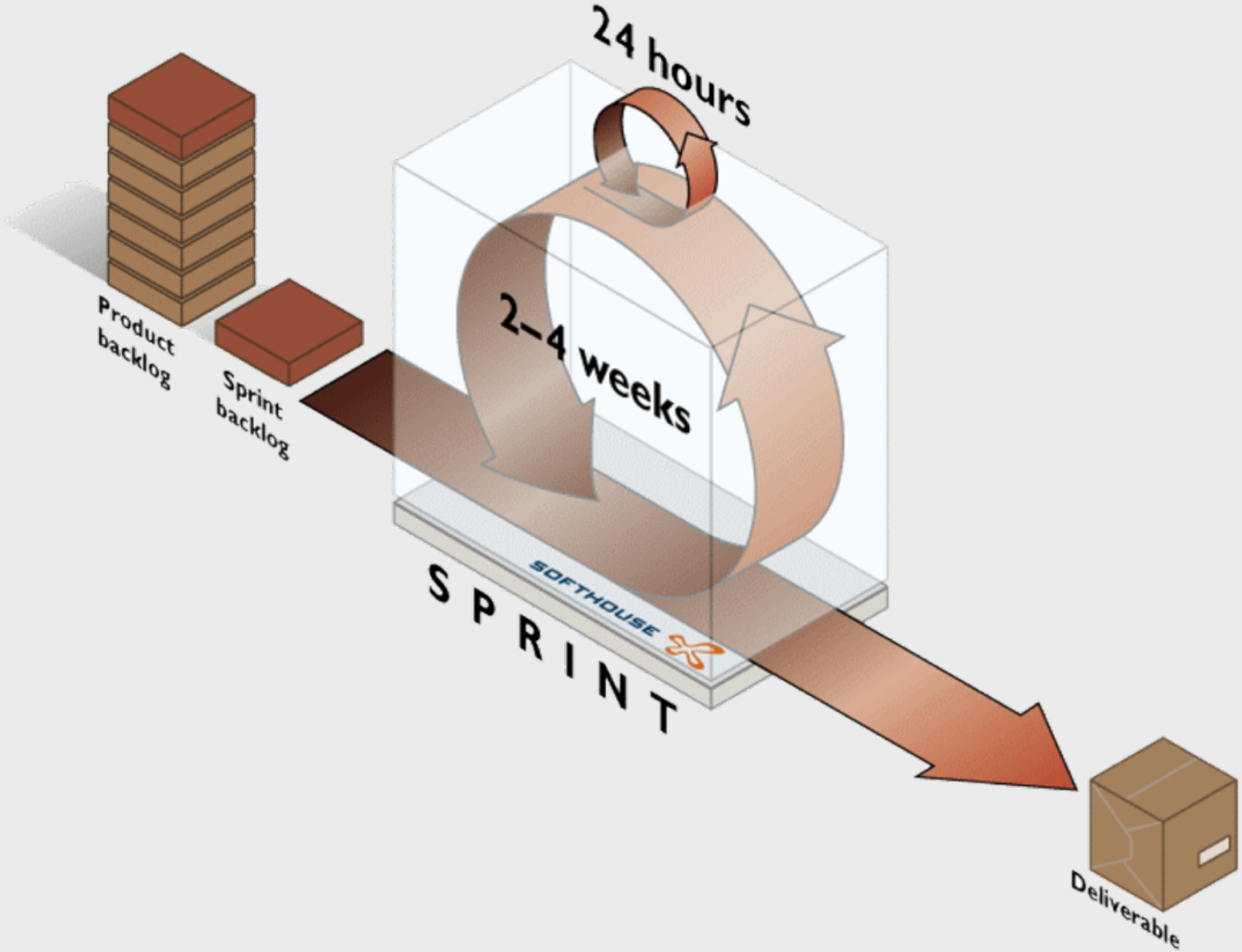
Scrum is a simple framework used to organize teams and get work done more productively with higher quality. It allows teams to choose the amount of work to be done and decide how best to do it, thereby providing a more enjoyable and productive working environment. Scrum focuses on prioritizing work based on business value, improving the usefulness of what is delivered, and increasing revenue, particularly early revenue.

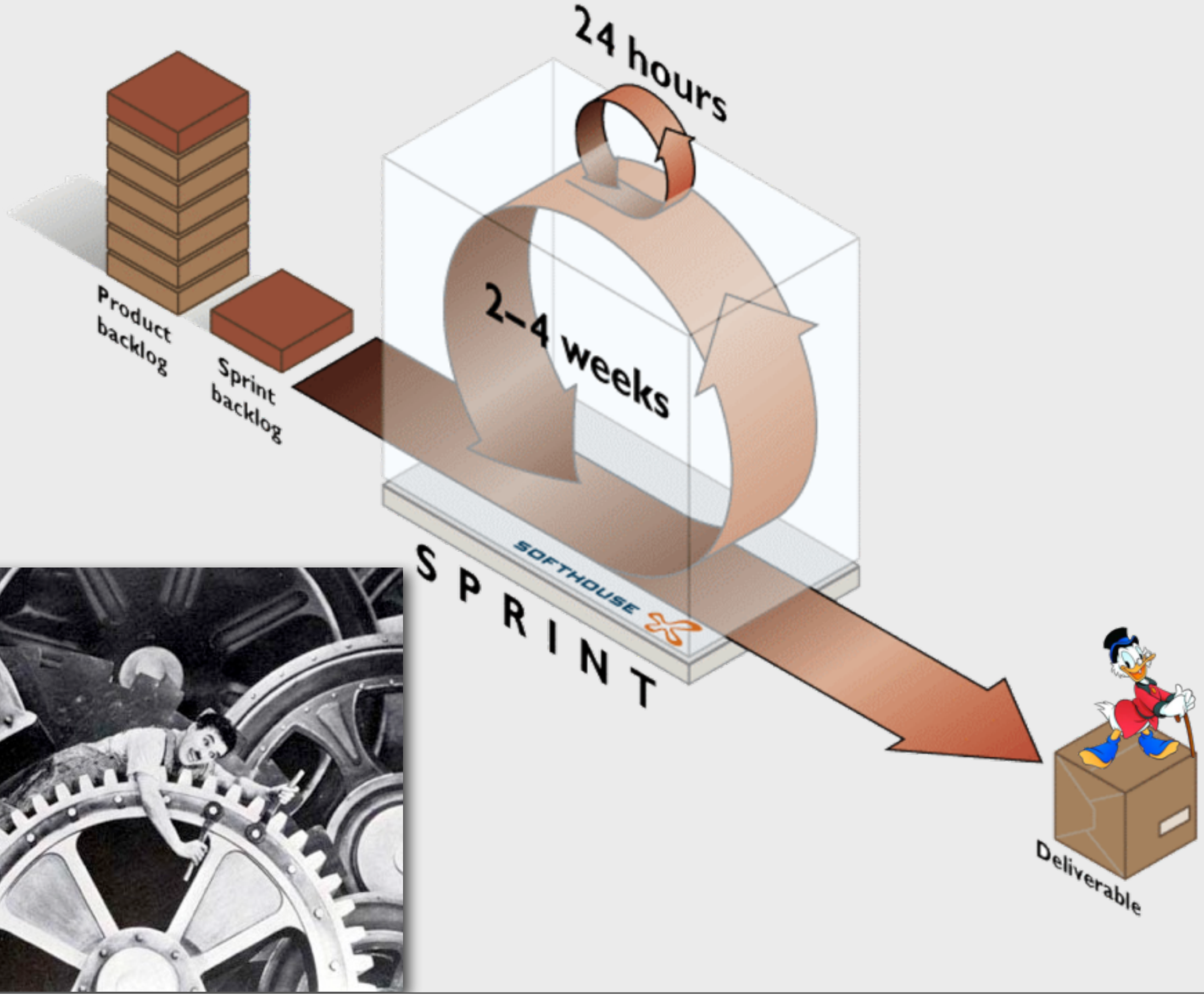
–Jeff Sutherland, *A Brief Introduction to Scrum*, 2007

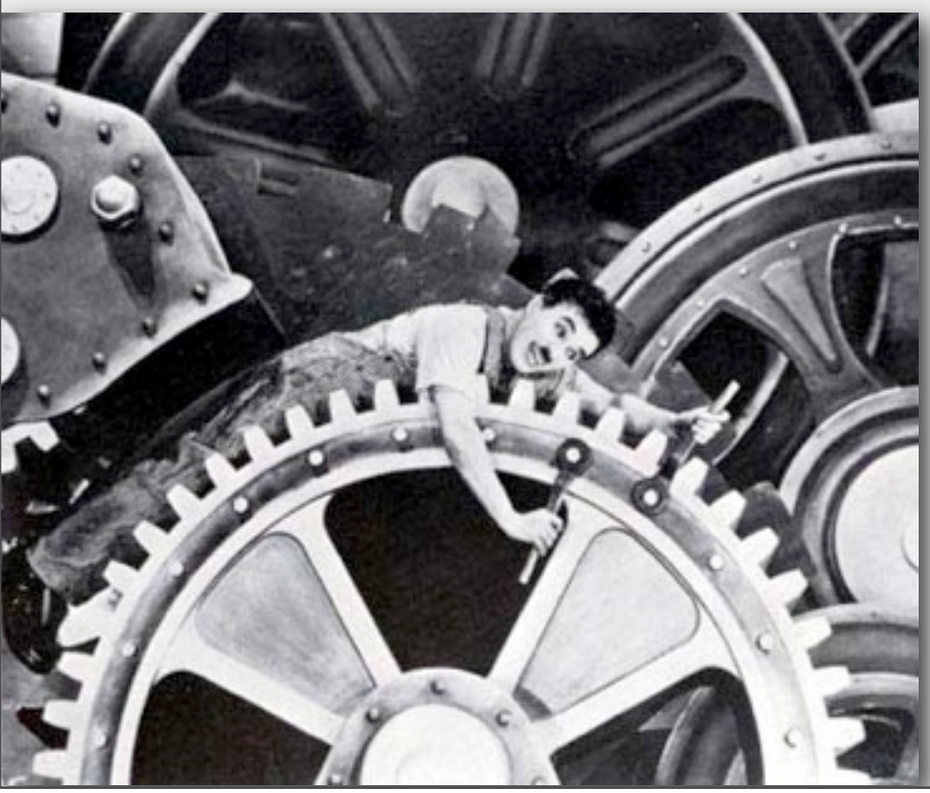
You Think Scrum is Any Better?

Designed to adapt to changing requirements during the development process at short, regular intervals, Scrum allows teams to prioritize customer requirements and adapt the work product in real time to customer needs. By doing this, Scrum provides what the customer wants at the time of delivery (improving customer satisfaction) while eliminating waste (work that is not highly valued by the customer).

–Jeff Sutherland, *A Brief Introduction to Scrum, 2007*


24 hours
2–4 weeks
Product backlog
Sprint backlog
SOFTHOUSE
SPRINT
Deliverable


24 hours
2–4 weeks
Product backlog
Sprint backlog
SOFTHOUSE
SPRINT
Deliverable



But merely extending knowledge a step further is not developing science. Breeding homing pigeons that could cover a given space with ever increasing rapidity did not give us the laws of telegraphy, nor did breeding faster horses bring us the steam locomotive.

–Edward J. v. K. Menge, *The Quarterly Review of Biology*, 1930

Our highest priority is to satisfy the customer through early and continuous delivery of valuable software.

–Agile Manifesto, *Principles behind the Agile Manifesto*

Who Is My Customer?

Our highest priority is to satisfy the customer through early and continuous delivery of valuable software.

–Agile Manifesto, *Principles behind the Agile Manifesto*

Friday, October 18, 13

- IBM pays me (my customer?)
- DARPA pays them (my customer?)
- They asked me to do science (my customer!)

- Computational Modeling of User Dynamic Behavior
- Computational Models of Trust and Social Capital
- Information Morphing Modeling
- Persuasiveness of Memes
- The Observability of Social Systems
- Culture-Dependent Social Media Modeling
- Dynamics of Influence in Social Networks
- Understanding the Optimal Immunization Policy
- Modeling and Identification of Campaign Target Audience
- Modeling and Predicting Competing Memes

- Real-Time and Large-Scale Social Media Mining
- Role and Function Discovery
- Detecting Malicious Users and Malware Propagation
- Emergent Topic Detection and Tracking
- Detecting Evolution History and Authenticity of Multimedia Memes
- Synchronistic Social Media Information and Social Proof Opinion Mining
- Community Detection and Tracking
- Interplay Across Multiple-Networks

- Crowd-sourcing Evidence Gathering to Formulate Counter-messaging Objectives
- Delivery and Evaluation of a Counter-messaging Campaign
- Optimal Target People Selection
- Automated Generation of Counter Messaging
- User Interfaces for Semi-Automatic Counter Messaging
- Controlling the Dynamics of Influence in Social Networks
- Influencing the Outcome of Competing Memes and Counter Messaging

Delivery and Evaluation of Counter-Messaging Campaign

- **Objective**: Deliver counter-messages and evaluate the effectiveness of counter-messaging campaign
- **Task Goals**:
  - Understand and Operationalize Methods of Delivering Counter-Messages
  - Accurate Evaluation of Affect of Campaign

Delivery and Evaluation of Counter-Messaging Campaign

- **Proposed Work:**
  - Model and Develop Set of Delivery Techniques
  - Take into account objectives, likelihood of accepting message, required incentives, etc.
- Method to Choose Delivery Technique
  - Public vs. Private, Choice of Communication Type, Etc.
  - Develop a Set of Methods & Metrics to Evaluate Impact of Messaging
  - Are objectives accomplished? What are the effects? What remains to be done?

Automated Generation of Counter Messaging

my part

- **Objective:** Automatically generate message content for counter-messaging campaign
- **Task Goals:**
  - Generate message content that includes influencing operation and considers features of message recipient(s)

Automated Generation of Counter Messaging

my part

- **Proposed Work:**
  - Model and Develop a Set of Seed Counter-Messaging Templates
  - Each template defines a message type to be generated
    - For example, a rumor refutation template might focus on exposing information inconsistencies
  - Automatic Retrieval of Templates Based on Situation
  - Composition of Relevant Templates
  - Instance-based learning
  - Optimization-based approaches

Twitter

- BTW, we work mostly with Tweets

Friday, October 18, 13

What I Am Supposed to Program

- Figure out the rhetorical strategy
- Consider who the target is
  - and then to whom to direct the messages
- Figure out the diction level
- Decide whom to sound like
- Figure out what other influences to embody
- Figure out what otherwise irrelevant facts to include

Oh, and...

- What is the personality of the addressee?
- What sentiments has this person displayed recently?
- Are there words or phrases or ideas that can be emphasized using poetic techniques?
- Are there any layered messages that can be embedded in word choice?

Oh, and...

- What is the personality of the addressee?
- What sentiments has this person displayed recently?
- Are there words or phrases or ideas that can be emphasized using poetic techniques?
- Are there any layered messages that can be embedded in word choice?

Friends, Romans, countrymen—lénd mé yóur éars

Oh, and...

- What is the personality of the addressee?
- What sentiments has this person displayed recently?
- Are there words or phrases or ideas that can be emphasized using poetic techniques?
- Are there any layered messages that can be embedded in word choice?

Oh, and...

- What is the personality of the addressee?
- What sentiments has this person displayed recently?
- Are there words or phrases or ideas that can be emphasized using poetic techniques?
- Are there any layered messages that can be embedded in word choice?

A dark man robbed me.

Oh, and...

- What is the personality of the addressee?
- What sentiments has this person displayed recently?
- Are there words or phrases or ideas that can be emphasized using poetic techniques?
- Are there any layered messages that can be embedded in word choice?

And Then There's

Beautiful Writing

...deep inside, we never quite forget the needs with which we were born: to be accepted as we are, without regard to our deeds; to be loved through the medium of our body; to be enclosed in another's arms; to occasion delight with the smell of our skin—all of these needs inspiring our relentless and passionately idealistic quest for someone to kiss and sleep with....

–Alain de Botton, *How to Think More About Sex, 2012*

My Goal

Something Between These Two

Hey, you know that the email apparently from Paypal is a phishing scam. You can read about it here: http://purportal.com/spam/2528/.

Also, note that the update link points here: http://maserverbn.com/ *and not to* http://paypal.com.

My Goal
Something Between These Two

The journey will be difficult. The road will be long. I face this challenge—I face this challenge with profound humility and knowledge of my own limitations, bút Í wíll also fáce ít with limitless faith in the capacity of the American people.

–Barrack Obama, *acceptance speech, Democratic Convention, 2008*

Interviewer: *How much rewriting do you do?*

Hemingway: *It depends. I rewrote the ending of* Farewell to Arms, *the last page of it, 39 times before I was satisfied.*

Interviewer: *Was there some technical problem there? What was it that had stumped you?*

Hemingway: *Getting the words right.*

nature & my code,
not a product owner,
stare at me every day

Nature, Science, and Understanding

my collaborators

my adversaries

Friday, October 18, 13

Rhetoric

- Blah
- Blah
- Blah

Ethos / Decorum

- Fitting in
- Meeting expectations
- Looking the right way to the audience
- Sounding the right way
- Writing the right way—for the audience

Make Them Listen

- Make them receptive
- Make them like & trust you
- Exhibit Virtue / Shared Values
- Exhibit Practical Wisdom / Street Smarts
- Show Disinterest / No skin in the game / Same skin in the game
- Have someone brag for you

Speak Like Them

- Code Grooming* (or Dog Whistles)
- Don't always speak in rational sentences
- Repeat codewords
- Find words that mean the opposite of your opponent's and negate them if you have to so your words are heard too: "*I think we are welcomed. But it was not a peaceful welcome.*"

–George W. Bush

* Like the way chimps groom to establish bonds

What Are The Ingredients?

- sentiment analysis
- personality assessment
- speech habits and personal corpora
- maybe an English parser (n-grams might suffice)

What Are The Ingredients?

- dictionaries
  - English dictionaries
  - syllabic dictionary
  - pronunciation dictionary
  - thesaurus (synonyms & antonyms)
  - emoticons, abbreviations
  - slang
  - idioms
  - rhyming dictionary
  - metaphor dictionary?

What Are The Ingredients?

- Corpora
  - ordinary articles and essays (representing different sentiments and personalities)
  - poetry
  - business tracts
  - religious tracts
  - tweetish tracts
  - corpora gleaned from targeted individuals (who might be people in the main target's social circles)
  - big pile of n-grams

What Are The Ingredients?

- A raft of poetic craft elements, analyzed in an Alexandrian setting:
  - meter
  - noise of the poem
  - rhyme
  - repetitions and echoes
  - line / sentence beginnings and endings
  - roughness
  - gradients
  - contrast
  - levels of scale
  - local symmetries
  - stillness...

What Are The Ingredients?

- Main argument, expressed pseudo-linguistically
- Contexts and modifiers for the main players in the sentences, which will result in subordinate clauses, adjectives, adverbs, and supporting sentences
- Other things “I” said before and things the target has said before (in case “I” need to fill in material for sonic or poetic affect)

Semantics?

- Nah, I think there will be only a network of sentences and phrases
- Soft / fuzzy / statistical matching
- Sourcing sentence / phrase / metaphor templates from appropriate corpora to establish diction levels and affinities
- Using machine optimization e.g. simulated annealing or a genetic algorithm

Semantics?

- Prefer particular words to establish sentiment and hence personality
- Use n-grams for correct grammar, and maybe a simple parser
- Being grammatical is not a strict requirement
- Otherwise, you got me exactly how I will do it

Sentiment Analysis

Journal of Language and Social Psychology
29(1) 24–54

DOI: 10.1177/0261927X09351676
http://jls.sagepub.com
SAGE

The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods

Yla R. Tausczik[1] and James W. Pennebaker[1]

Linguistic Inquiry and Word Count

1	All pronouns	18	Anger	35	Family	52	Home
2	1st person singular	19	Sadness	36	Humans	53	Sport/exercise
3	1st person plural	20	Cognition	37	Time	54	TV/movies
4	Total 1st person	21	Cause@Causation	38	Past	55	Music
5	Total 2nd person	22	Insight	39	Present	56	Money
6	Total 3rd person	23	Discrepancy	40	Future	57	Metaphysical
7	Negations	24	Inhibition	41	Space	58	Religion
8	Assents	25	Tentativeness	42	Up	59	Death
9	Articles	26	Certainty	43	Down	60	Physical states/factors
10	Prepositions	27	Sensation/perception	44	Inclusion	61	Symptoms & sensations
11	Numbers	28	Seeing	45	Exclusion	62	Sexual
12	Affect	29	Hearing	46	Motion	63	Eating/drinking
13	Positive affect	30	Touching	47	Occupation	64	Sleeping/dreaming
14	Positive feelings	31	Social	48	School	65	Grooming
15	Optimism	32	Communication	49	Job	66	Swear words
16	Negative affect	33	Reference to others	50	Achievement	67	Non-fluencies
17	anxiety	34	Friends	51	Leisure	68	Fillers

contain*	20	24					
contented*	12	13					
continu*	37						
contradic*	12	16	18	20	24	31	32
control*	12	13	15	20	24	47	50
convers*	31	32					
convinc*	12	13	15				
cook*	60	63					

convers*	Social	Communication	
convinc*	Affect	Positive Affect	Optimism
cook*	Physical States	Eating / Drinking	

Hemingway: All Stories

Total words [expanded words] (talkativeness, verbal fluency): 114063 [114063]
Different words: 9276
Average word length: 4.0
Words longer than 6 letters (education, social class): 13148 (11.5%)
Unique words longer than 6 letters: 4078 (44.0%)
Number of sentences (approx.): 9455
Average sentence length (verbal fluency, cognitive complexity): 12.1
Words captured (informal, nontechnical language): 82743 (72.5%)

Gabriel: Patterns of Software

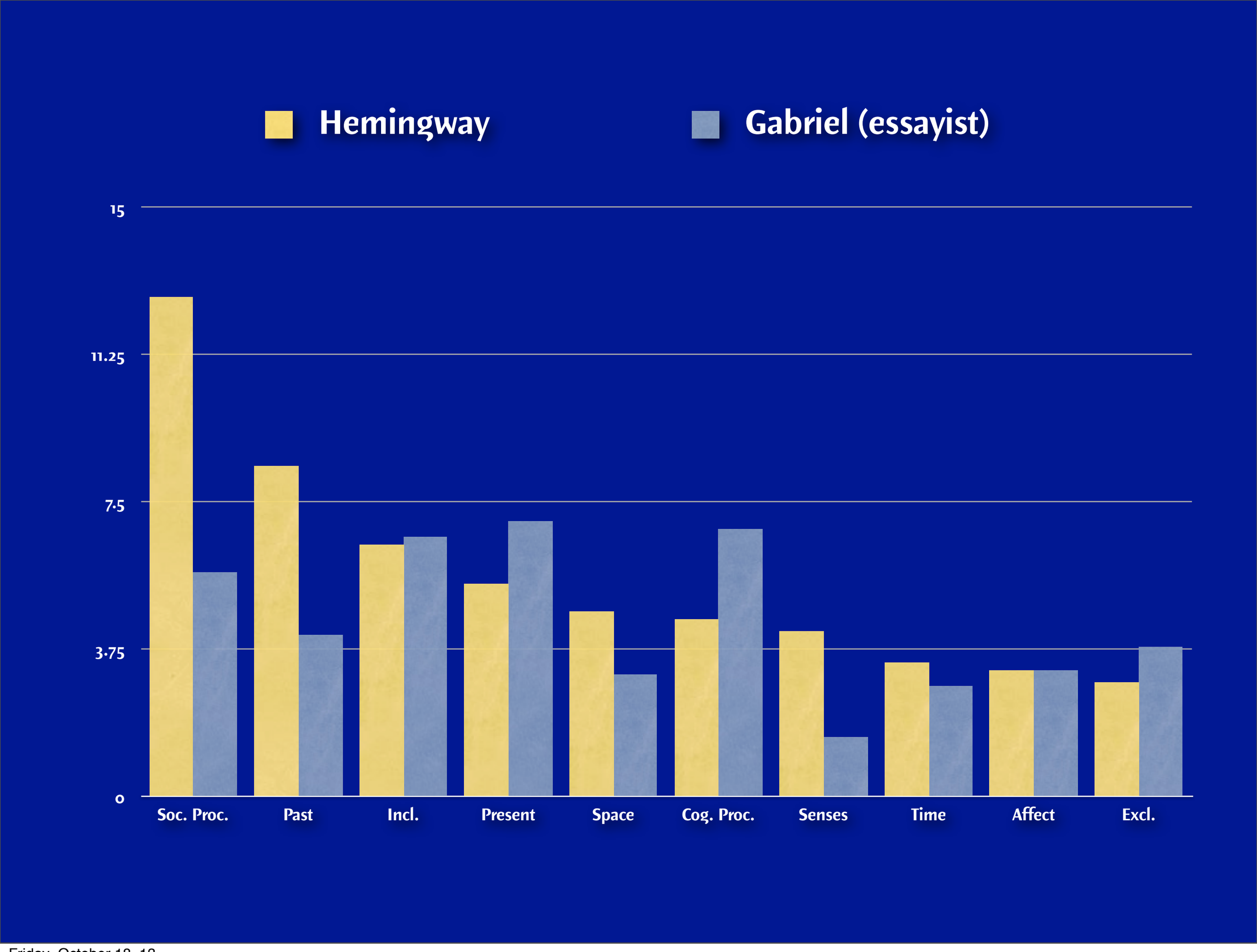
Total words [expanded words] (talkativeness, verbal fluency): 94306 [94306]
Different words: 9926
Average word length: 4.7
Words longer than 6 letters (education, social class): 22458 (23.8%)
Unique words longer than 6 letters: 5882 (59.3%)
Number of sentences (approx.): 4369
Average sentence length (verbal fluency, cognitive complexity): 21.6
Words captured (informal, nontechnical language): 63359 (67.2%)

Hemingway

1: SOCIAL_PROCESSES 14519 12.7%
2: PAST_TENSE_VB 9567 8.4%
3: INCLUSIVE 7317 6.4%
4: PRESENT_TENSE_VB 6145 5.4%
5: SPACE 5334 4.7%
6: COGNITIVE_PROCESSES 5174 4.5%
7: SENSORY_PROCESSES 4767 4.2%
8: TIME 3829 3.4%
9: AFFECT 3675 3.2%
10: EXCLUSIVE 3253 2.9%
11: COMMUNICATION 2944 2.6%
12: HEARING 2692 2.4%
13: PHYSICAL_STATES 2426 2.1%
14: POSITIVE_EMOTIONS 2167 1.9%
15: DISCREPANCY 2158 1.9%
16: UP 1869 1.6%
17: MOTION 1857 1.6%
18: BODY_STATES 1616 1.4%
19: CERTAINTY 1557 1.4%
20: INSIGHT 1546 1.4%

Gabriel (essayist)

1: PRESENT_TENSE_VB 6615 7.0%
2: COGNITIVE_PROCESSES 6445 6.8%
3: INCLUSIVE 6239 6.6%
4: SOCIAL_PROCESSES 5332 5.7%
5: PAST_TENSE_VB 3842 4.1%
6: EXCLUSIVE 3621 3.8%
7: OCCUPATION 3415 3.6%
8: AFFECT 2986 3.2%
9: SPACE 2891 3.1%
10: TIME 2607 2.8%
11: POSITIVE_EMOTIONS 2160 2.3%
12: TENTATIVE 1971 2.1%
13: DISCREPANCY 1952 2.1%
14: INSIGHT 1909 2.0%
15: ACHIEVEMENT 1463 1.6%
16: SENSORY_PROCESSES 1402 1.5%
17: JOB/WORK 1297 1.4%
18: COMMUNICATION 1124 1.2%
19: CAUSATION 1119 1.2%
20: NEGATION 1001 1.1%

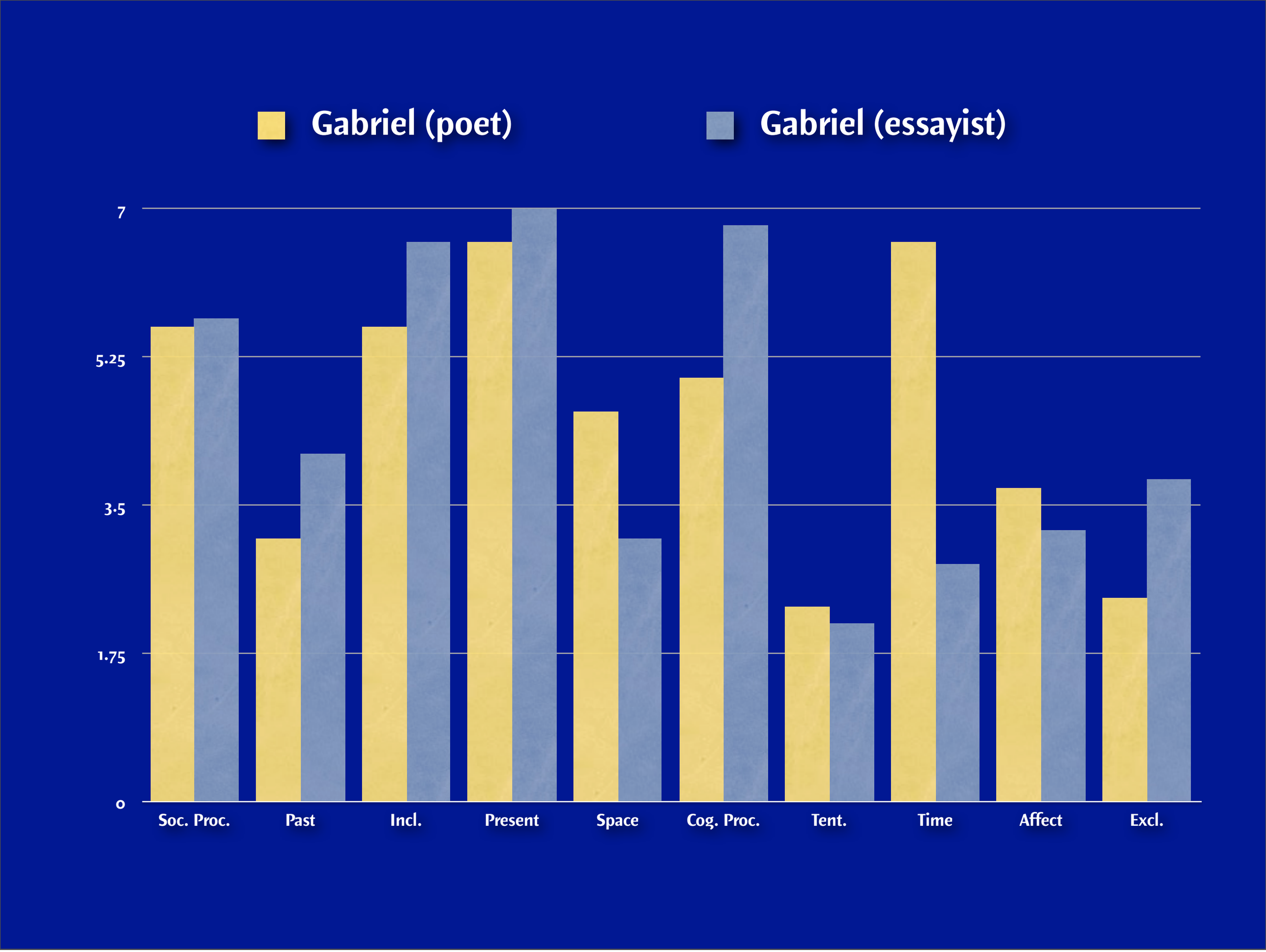

Hemingway
Gabriel (essayist)
15
11.25
7.5
3.75
0
Soc. Proc.
Past
Incl.
Present
Space
Cog. Proc.
Senses
Time
Affect
Excl.

Gabriel (poet)

1: TIME 1134 6.6%
2: PRESENT_TENSE_VB 1120 6.6%
3: INCLUSIVE 959 5.6%
4: SOCIAL_PROCESSES 951 5.6%
5: COGNITIVE_PROCESSES 845 5.0%
6: SPACE 778 4.6%
7: AFFECT 625 3.7%
8: PAST_TENSE_VB 537 3.1%
9: EXCLUSIVE 407 2.4%
10: TENTATIVE 394 2.3%
11: POSITIVE_EMOTIONS 391 2.3%
12: DISCREPANCY 341 2.0%
13: SENSORY_PROCESSES 333 2.0%
14: PHYSICAL_STATES 291 1.7%
15: UP 253 1.5%
16: NEGATIVE_EMOTIONS 237 1.4%
17: INSIGHT 231 1.4%
18: NEGATION 213 1.2%
19: CERTAINTY 194 1.1%
20: LEISURE 184 1.1%

Gabriel (essaysist)

1: PRESENT_TENSE_VB 6615 7.0%
2: COGNITIVE_PROCESSES 6445 6.8%
3: INCLUSIVE 6239 6.6%
4: SOCIAL_PROCESSES 5332 5.7%
5: PAST_TENSE_VB 3842 4.1%
6: EXCLUSIVE 3621 3.8%
7: OCCUPATION 3415 3.6%
8: AFFECT 2986 3.2%
9: SPACE 2891 3.1%
10: TIME 2607 2.8%
11: POSITIVE_EMOTIONS 2160 2.3%
12: TENTATIVE 1971 2.1%
13: DISCREPANCY 1952 2.1%
14: INSIGHT 1909 2.0%
15: ACHIEVEMENT 1463 1.6%
16: SENSORY_PROCESSES 1402 1.5%
17: JOB/WORK 1297 1.4%
18: COMMUNICATION 1124 1.2%
19: CAUSATION 1119 1.2%
20: NEGATION 1001 1.1%


Gabriel (poet)
Gabriel (essayist)
7
5.25
3.5
1.75
0
Soc. Proc.
Past
Incl.
Present
Space
Cog. Proc.
Tent.
Time
Affect
Excl.

Personality: Big Five

- Extraversion vs. Introversion
- Emotional stability vs. Neuroticism
- Agreeableness vs. Disagreeable
- Conscientiousness vs. Unconscientious
- Openness to experience

Published in final edited form as:
J Res Pers. 2010 June 1; 44(3): 363–373. doi:10.1016/j.jrp.2010.04.001.

Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers

Tal Yarkoni
University of Colorado at Boulder

Abstract

Previous studies have found systematic associations between personality and individual differences in word use. Such studies have typically focused on broad associations between major personality domains and aggregate word categories, potentially masking more specific associations. Here I report the results of a large-scale analysis of personality and word use in a large sample of blogs (N=694). The size of the dataset enabled pervasive correlations with personality to be identified for a broad range of lexical variables, including both aggregate word categories and individual English words. The results replicated category-level findings from previous offline studies, identified numerous novel associations at both a categorical and single-word level, and underscored the value of complementary approaches to the study of personality and word use.

People differ considerably from each other in their habitual patterns of thought, feeling and action. Not surprisingly, these differences are reflected not only in what people think, feel, and do, but also in what they *say* about what they think, feel, or do. Recent studies have identified systematic associations between personality and language use in a variety of different contexts,

Published in final edited form as:
J Res Pers. 2010 June 1; 44(3): 363–373. doi:10.1016/j.jrp.2010.04.001.

Personality in 100,000 Words: A large-scale analysis of personality and word use among bloggers

Tal Yarkoni
University of Colorado at Boulder

Abstract

Previous studies have found systematic associations between personality and individual differences in word use. Such studies have typically focused on broad associations between major personality domains and aggregate word categories, potentially masking more specific associations. Here I report the results of a large-scale analysis of personality and word use in a large sample of blogs (N=694). The size of the dataset enabled pervasive correlations with personality to be identified for a broad range of lexical variables, including both aggregate word categories and individual English words. The results replicated category-level findings from previous offline studies, identified numerous novel associations at both a categorical and single-word level, and underscored the value of complementary approaches to the study of personality and word use.

People differ considerably from each other in their habitual patterns of thought, feeling and action. Not surprisingly, these differences are reflected not only in what people think, feel, and do, but also in what they *say* about what they think, feel, or do. Recent studies have identified systematic associations between personality and language use in a variety of different contexts,

Personality

- Linear combination of LIWC scores

most

Big Five
(provisional example)

least

Hemingway	
Conscientiousness	2.04
Extraversion	10.86
Openness	-23.95
Agreeableness	38.05
Neuroticism	-4.53

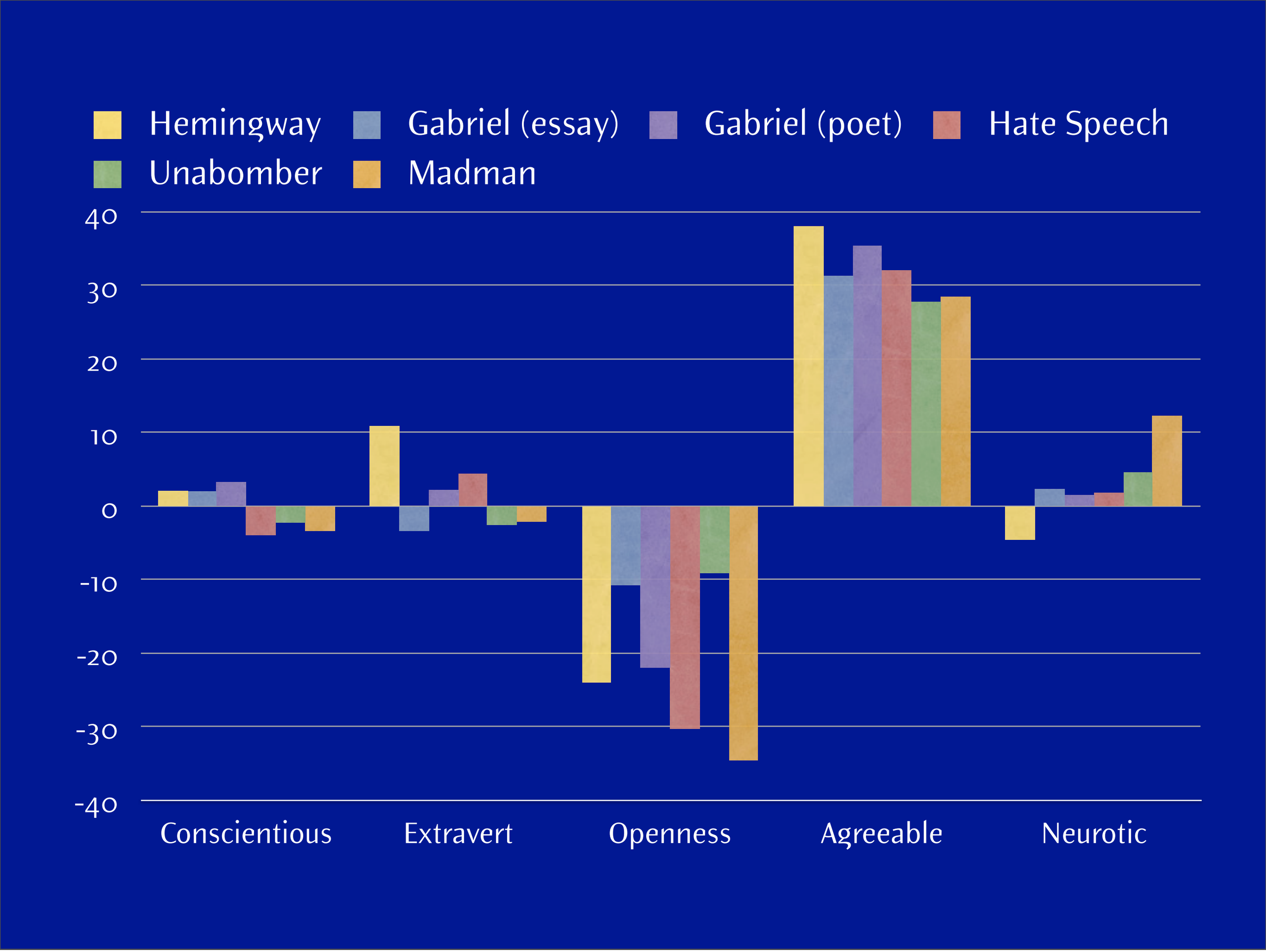
Gabriel (essay)	
Conscientiousness	1.99
Extraversion	-3.32
Openness	-10.73
Agreeableness	31.33
Neuroticism	2.33

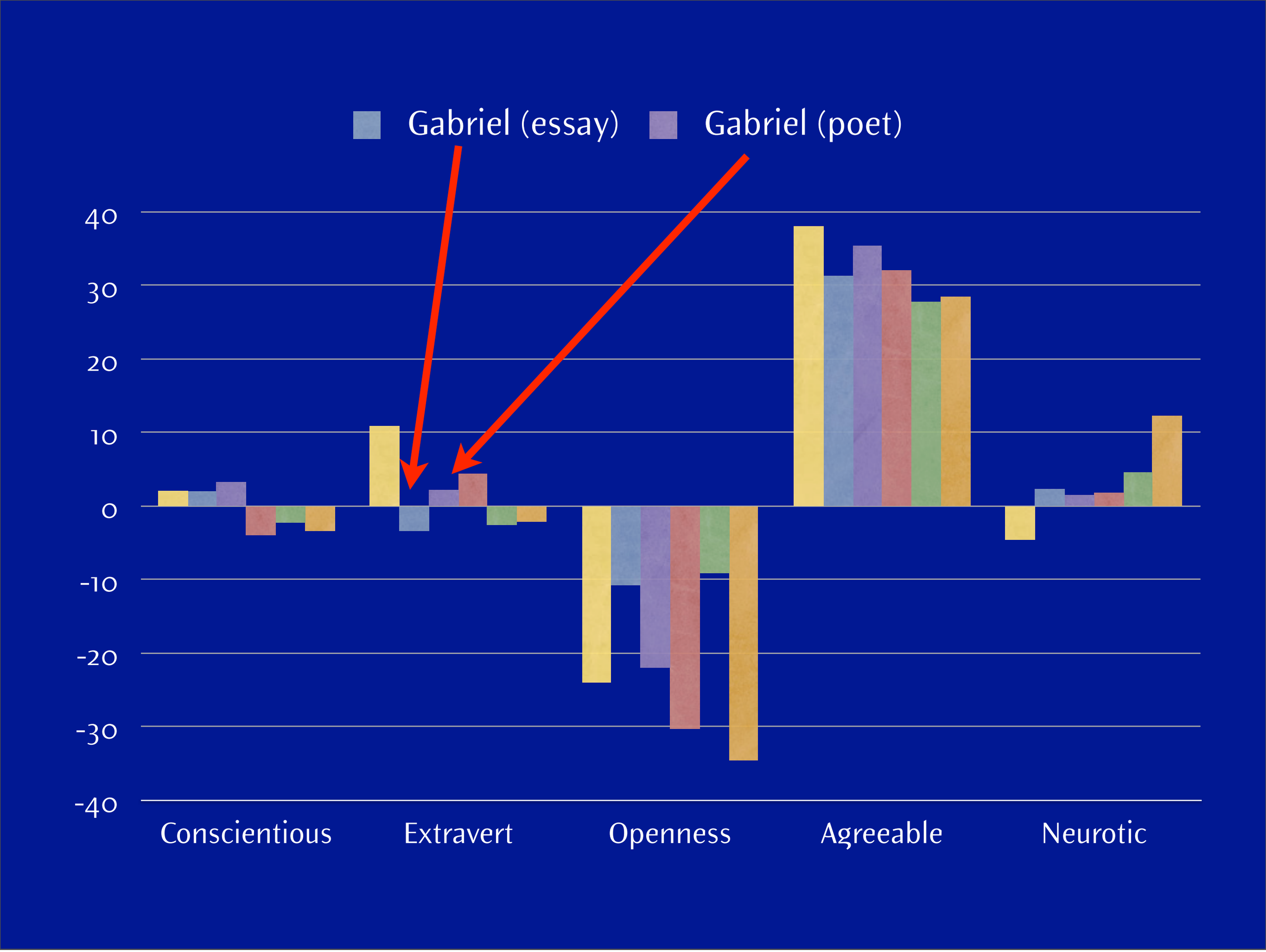
Hate Speech	
Conscientiousness	-3.92
Extraversion	4.37
Openness	-30.27
Agreeableness	32.05
Neuroticism	1.81

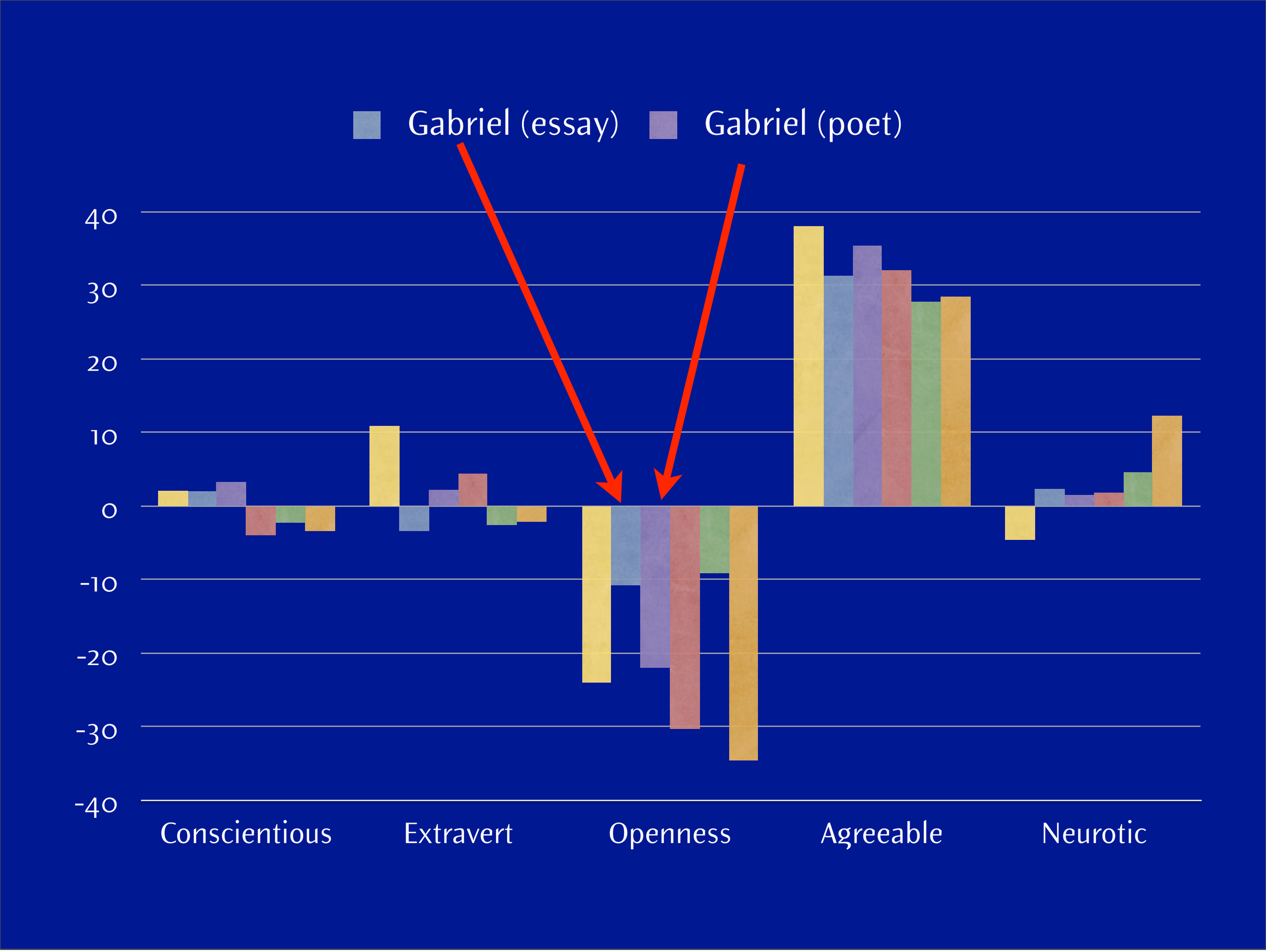
Unabomber	
Conscientiousness	-2.23
Extraversion	-2.56
Openness	-9.06
Agreeableness	27.76
Neuroticism	4.56

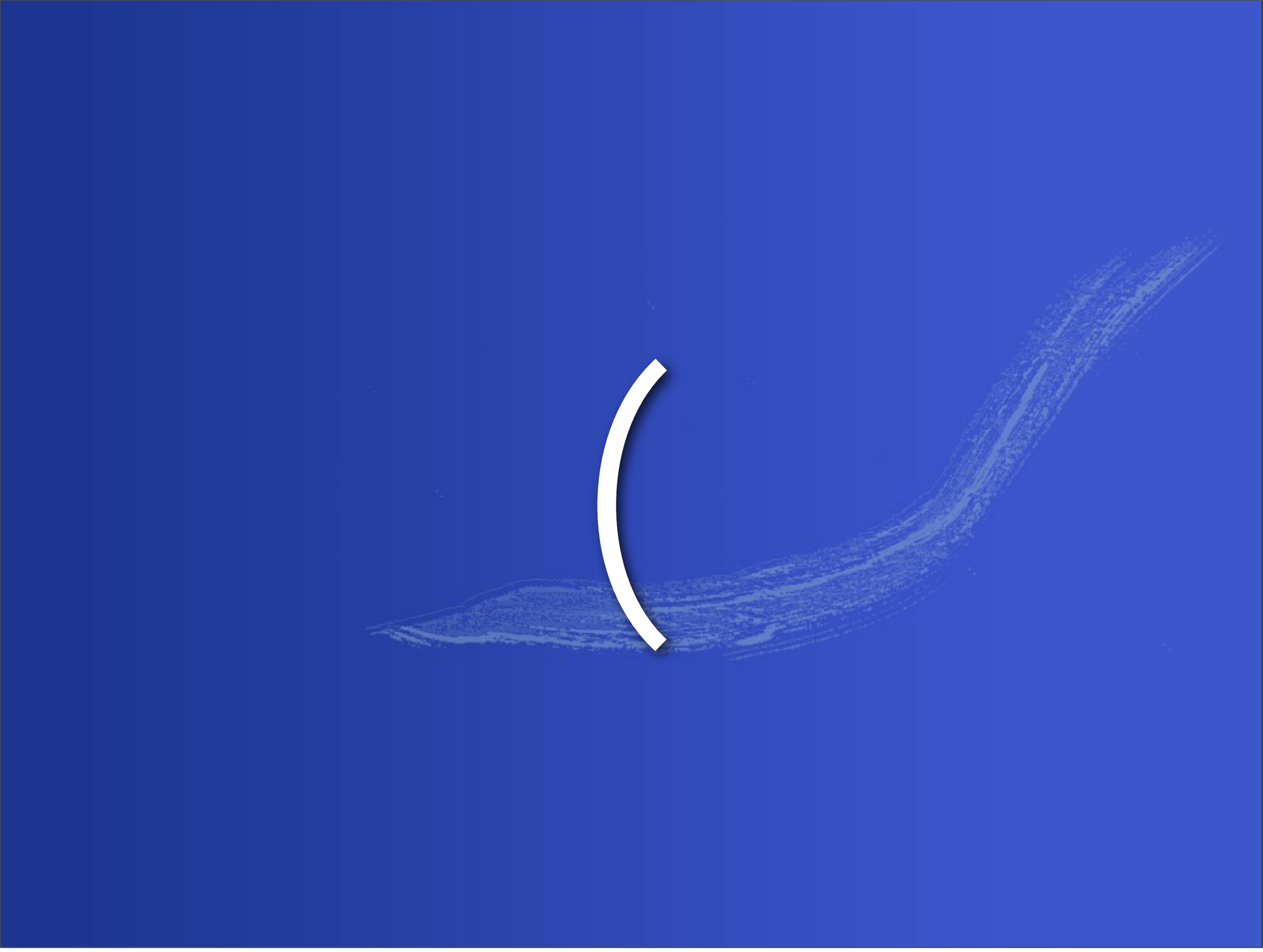
Gabriel (poet)	
Conscientiousness	3.24
Extraversion	2.19
Openness	-21.96
Agreeableness	35.38
Neuroticism	1.48

CS person gone mad	
Conscientiousness	-3.36
Extraversion	-2.08
Openness	-34.53
Agreeableness	28.48
Neuroticism	12.26


Hemingway
Gabriel (essay)
Gabriel (poet)
Hate Speech
Unabomber
Madman
40
30
20
10
0
-10
-20
-30
-40
Conscientious
Extravert
Openness
Agreeable
Neurotic


Gabriel (essay)
Gabriel (poet)
40
30
20
10
0
-10
-20
-30
-40
Conscientious
Extravert
Openness
Agreeable
Neurotic


Gabriel (essay)
Gabriel (poet)
40
30
20
10
0
-10
-20
-30
-40
Conscientious
Extravert
Openness
Agreeable
Neurotic



LIWC apparently tracks genre
(a little)

what if...

- Instead of these traits:
  - Conscientiousness
  - Agreeableness
  - Openness
  - Extraversion
  - Neuroticism
- We use these traits:
  - Poetry
  - Fiction
  - Nonfiction

Training Targets

Text File	Poetry	Fiction	Nonfiction
Poemsrpg (P)	85.0	-10.0	-10.0
Leaves of Grass (P)	95.0	-30.0	-50.0
Traditional Salvation (F)	-10.0	80.0	-25.0
Hemingway (F)	-10.0	95.0	-75.0
Patterns Of Software (NF)	-35.0	-5.0	95.0
Writers' Workshop (NF)	-10.0	-2.0	90.0
Faulkner (F)	-5.0	95.0	-65.0
Ulysses (F)	-5.0	90.0	-15.0
Emily Dickinson (P)	95.0	-25.0	-80.0
Unabomber (NF)	-70.0	-50.0	85.0
Wizard of Oz (F)	-25.0	85.0	-35.0
Call Of The Wild (F)	-12.0	87.0	-55.0
Huckleberry Finn (F)	-5.0	45.0	-40.0
Metamorphosis (F)	-25.0	70.0	-35.0
Origin Of Species (NF)	-80.0	-10.0	75.0

Training Files & Classifications

Text File	Genre
Poemsrpg (P)	Poetry
Leaves of Grass (P)	Poetry
Traditional Salvation (F)	Fiction
Hemingway (F)	Fiction
Patterns Of Software (NF)	Nonfiction
Writers' Workshop (NF)	Nonfiction
Faulkner (F)	Fiction
Ulysses (F)	Fiction
Emily Dickinson (P)	Poetry
Unabomber (NF)	Nonfiction
Wizard of Oz (F)	Fiction
Call Of The Wild (F)	Fiction
Huckleberry Finn (F)	Fiction
Metamorphosis (F)	Fiction
Origin Of Species (NF)	Nonfiction

Text File	Genre	Poetry	Fiction	Nonfiction
Knott (P)	Poetry			
Trakl (P)	Poetry			
Lanier (P)	Poetry			
The Wasteland (P)	Poetry			
Moby Dick (F)	Fiction			
Gay Stories (F)	Fiction			
To Kill a Mockingbird (F)	Fiction			
Hamlet (?)	Fiction[Poetry]	-5.14	24.7	-24.49
Bertrand Russell (NF)	Nonfiction			
Charles Babbage (NF)	Nonfiction			
Darwin (NF)	Nonfiction			
Crazy CS Person (NF)	Poetry	-0.12	13.59	-10.94
Bible (?)	Fiction[Poetry]	-9.6	42.30	-37.75
Pete Turchi's New Book (NF)	Fiction[Nonfiction]	-22.38	16.83	-5.14

$-5.0 \leq P$ (poetry)
$-5.0 \leq NF$ (nonfiction)
otherwise (fiction)

$-10.0 \leq P < -5.0$ (poetry mixin)
$-10.0 \leq NF < -5.0$ (nonfiction mixin)
$35.0 \leq F$ (fiction mixin)

Text File	Genre	Poetry	Fiction	Nonfiction
Gribble / Fedora (P)	Poetry			
Janet Holmes / Humanophone (P)	Fiction[Poetry]	-6.38	35.6	-27.39
Janet Holmes / F2F (P)	Poetry			
Front Page NYT Article (NF)	Fiction[Nonfiction]	-13.45	24.59	-7.73
Richard Schmitt / Kodiak (F)	Poetry[Fiction]	-4.73	46.18	-33.03
Richard Schmitt / A Year of Counseling (F)	Poetry[Fiction]	-4.33	36.07	-31.33
Harper / Prac. Found. for Prog. Lang (NF)	Nonfiction			
Ellen Bryant Voigt / Song and Story (P)	Poetry			
Tennyson / In Memoriam (P)	Poetry			
US Constitution (NF)	Nonfiction			
Tom Lux / I Love You Sweatheart (P)	Fiction	-12.44	40.13	-35.43
rpg / Sharp Tone (P)	Poetry			
Cass Pursell / Men and Stones (F)	Fiction			
Proust's Longest Sentence (F)	Fiction			

$-5.0 \leq$ P (poetry)
$-5.0 \leq$ NF (nonfiction)
otherwise (fiction)

$-10.0 \leq$ P < -5.0 (poetry mixin)
$-10.0 \leq$ NF < -5.0 (nonfiction mixin)
$35.0 \leq$ F (fiction mixin)

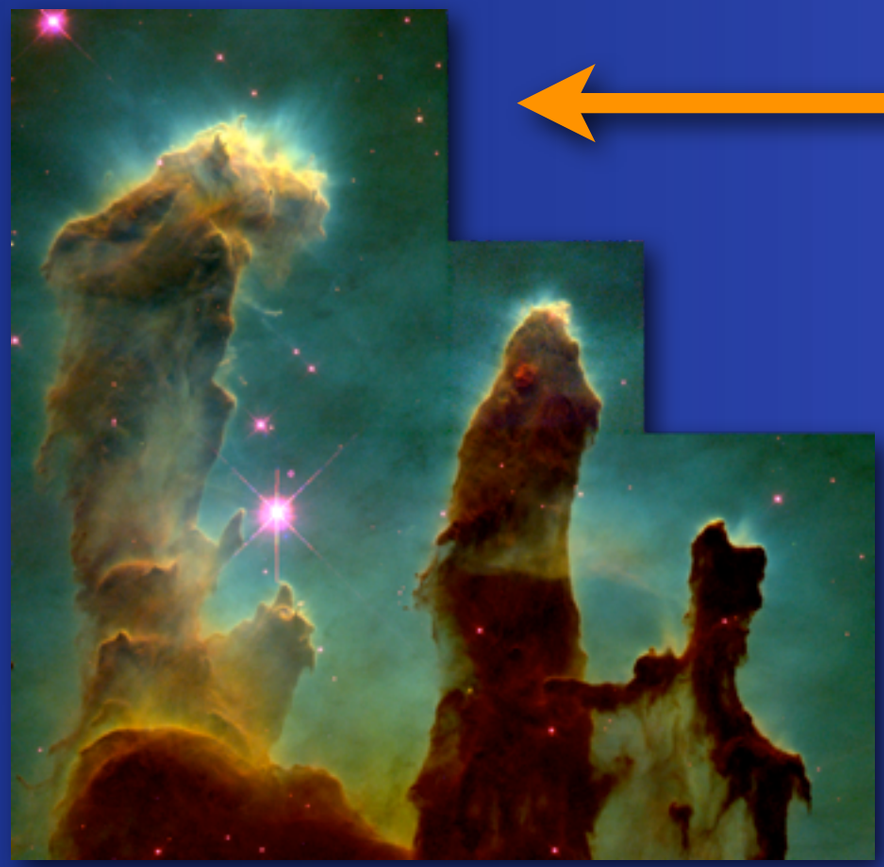
Surprising Observation

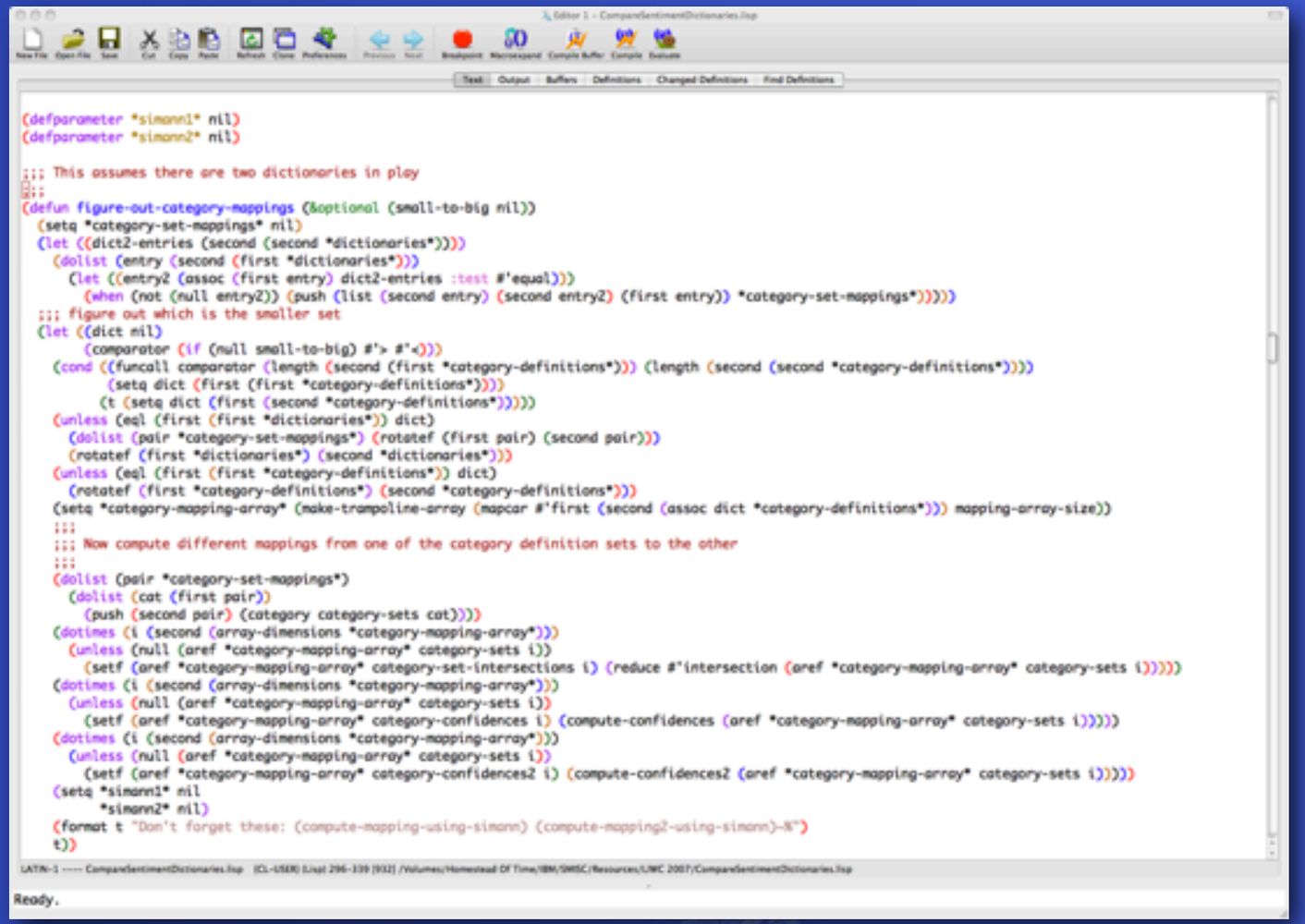
- Fiction is not special
- That is, everything looks like fiction—at least a little

-5.0 ≤ P (poetry)
-5.0 ≤ NF (nonfiction)
otherwise (fiction)

-10.0 ≤ P < -5.0 (poetry mixin)
-10.0 ≤ NF < -5.0 (nonfiction mixin)
35.0 ≤ F (fiction mixin)

Friday, October 18, 13





```
(defparameter *simann1* nil)
(defparameter *simann2* nil)

;;; This assumes there are two dictionaries in play
;;;
(defun figure-out-category-mappings (&optional (small-to-big nil))
  (setq *category-set-mappings* nil)
  (let ((dict2-entries (second (second *dictionaries*))))
    (dolist (entry (second (first *dictionaries*)))
      (let ((entry2 (assoc (first entry) dict2-entries :test #'equal)))
        (when (not (null entry2)) (push (list (second entry) (second entry2) (first entry)) *category-set-mappings*)))))
  ;;; figure out which is the smaller set
  (let ((dict nil)
        (comparator (if (null small-to-big) #'> #'<)))
    (cond ((funcall comparator (length (second (first *category-definitions*))) (length (second (second *category-definitions*))))
           (setq dict (first (first *category-definitions*)))
           (t (setq dict (first (second *category-definitions*)))))
    (unless (eql (first (first *dictionaries*)) dict)
      (dolist (pair *category-set-mappings*) (rotatef (first pair) (second pair)))
      (rotatef (first *dictionaries*) (second *dictionaries*)))
    (unless (eql (first (first *category-definitions*)) dict)
      (rotatef (first *category-definitions*) (second *category-definitions*)))
    (setq *category-mapping-array* (make-trampoline-array (mapcar #'first (second (assoc dict *category-definitions*))) mapping-array-size))
    ;;;
    ;;; Now compute different mappings from one of the category definition sets to the other
    ;;;
    (dolist (pair *category-set-mappings*)
      (dolist (cat (first pair))
        (push (second pair) (category category-sets cat))))
    (dotimes (i (second (array-dimensions *category-mapping-array*)))
      (unless (null (aref *category-mapping-array* category-sets i))
        (setf (aref *category-mapping-array* category-set-intersections i) (reduce #'intersection (aref *category-mapping-array* category-sets i)))))
    (dotimes (i (second (array-dimensions *category-mapping-array*)))
      (unless (null (aref *category-mapping-array* category-sets i))
        (setf (aref *category-mapping-array* category-confidences i) (compute-confidences (aref *category-mapping-array* category-sets i)))))
    (dotimes (i (second (array-dimensions *category-mapping-array*)))
      (unless (null (aref *category-mapping-array* category-sets i))
        (setf (aref *category-mapping-array* category-confidences2 i) (compute-confidences2 (aref *category-mapping-array* category-sets i)))))
    (setq *simann1* nil
          *simann2* nil)
    (format t "Don't forget these: (compute-mapping-using-simann) (compute-mapping2-using-simann)~%")
    t))
```

Ready.

$$\frac{}{\Phi;\Delta;\Gamma \vdash x \in \Gamma(x)}\ [\text{GT-Var}] \qquad \frac{\Phi;\Delta \vdash T\ ok \quad \Phi;\Delta;\Gamma \vdash e \in S}{\Phi;\Delta;\Gamma \vdash (T)e \in T}\ [\text{GT-Cast}]$$

$$\frac{\Phi \vdash T\ includes\ \text{init}(\overline{S}) \quad \Phi;\Delta;\Gamma \vdash \overline{e} \in \overline{S} \quad \Phi;\Delta \vdash T\ ok}{\Phi;\Delta;\Gamma \vdash \text{new}\ T(\overline{e}) \in T\ annotate\ [\overline{e} :: \overline{S}]}\ [\text{GT-New}]$$

$$\frac{fields(S) = \overline{T}\ \overline{f} \quad \Phi;\Delta;\Gamma \vdash e \in T \quad \Delta \vdash T <: S \quad \Delta \vdash P <: S\ \text{and}\ f_i \in fields(P)\ \text{implies}\ P = S}{\Phi;\Delta;\Gamma \vdash e.f_i \in T_i\ annotate\ [e :: S]}\ [\text{GT-Field}]$$

$$\frac{\Phi;\Delta \vdash \overline{T}\ ok \quad \Phi;\Delta;\Gamma \vdash e_0 \in T_0 \quad \Phi;\Delta;\Gamma \vdash \overline{e} \in \overline{S} \quad mtype(n; bound_\Delta(T_0)) = P.\langle \overline{X}\ \text{extends}\ \overline{S}\ \text{with}\ \overline{\P[\diamond} > S\ n(\overline{U}\ \overline{x}) \quad \Delta \vdash \overline{T} <: [\overline{X} \mapsto \overline{T}]\overline{S} \quad \Phi \vdash \overline{T}\ includes\ [\overline{X} \mapsto \overline{T}]\overline{\P[\diamond} \quad \Delta \vdash \overline{S} <: [\overline{X} \mapsto \overline{T}]\overline{U}}{\Phi;\Delta;\Gamma \vdash e_0.n\langle\overline{T}\rangle(\overline{e}) \in [\overline{X} \mapsto \overline{T}]S\ annotate\ [e_0 \in P]}\ [\text{GT-Invk}]$$

$$\frac{\Delta \vdash \overline{S} <: \overline{S} \quad \Phi;\Delta \vdash T\ ok \quad \Phi;\Delta \vdash \overline{S}\ ok \quad \Phi \vdash T\ includes\ \text{init}(\overline{S}) \quad \Phi;\Delta;\Gamma \vdash \overline{e} \in \overline{S}}{\Phi;\Delta;\Gamma \vdash \text{new}\ T(\overline{e} :: \overline{S}) \in T}\ [\text{GT-Ann-New}]$$

$$\frac{fields(S) = \overline{T}\ \overline{f} \quad \Phi;\Delta \vdash S\ ok \quad \Phi;\Delta;\Gamma \vdash e \in T \quad \Delta \vdash T <: S}{\Phi;\Delta;\Gamma \vdash [e :: S].f_i \in T_i}\ [\text{GT-Ann-Field}]$$

$$\frac{\Phi;\Delta \vdash \overline{T}\ ok \quad \Phi;\Delta;\Gamma \vdash e_0 \in T_0 \quad \Phi;\Delta;\Gamma \vdash \overline{e} \in \overline{S} \quad mtype(n; 0) = P.\langle \overline{X}\ \text{extends}\ \overline{S}\ \text{with}\ \overline{\P[\diamond} > S\ n(\overline{U}\ \overline{x}) \quad \Delta \vdash \overline{T} <: [\overline{X} \mapsto \overline{T}]\overline{S} \quad \Phi \vdash \overline{T}\ includes\ [\overline{X} \mapsto \overline{T}]\overline{\P[\diamond} \quad \Delta \vdash \overline{S} <: [\overline{X} \mapsto \overline{T}]\overline{U}}{\Phi;\Delta;\Gamma \vdash [e_0 \circ 0].n\langle\overline{T}\rangle(\overline{e}) \in [\overline{X} \mapsto \overline{T}]S}\ [\text{GT-Ann-Invk}]$$

I Don't Have

- A customer who knows what is required / desired
- Someone to interact with who can inform me what to do
- A boss with a mind that changes now and then

I Do Have

- Nature who never wavers but is generally mute
- The software I create which mediates my exploration of nature
- My own insight, which comes and goes but is more important than the actual code
- Mystery which with insight suggests changes as part of the process of exploration

Therefore

- Individuals and interactions over processes and tools
- Working software over comprehensive documentation
- Customer collaboration over contract negotiation
- Responding to change over following a plan

Therefore

- ~~Individuals and interactions over processes and tools~~
- ~~Working software over comprehensive documentation~~
- ~~Customer collaboration over contract negotiation~~
- ~~Responding to change over following a plan~~

Therefore

- Nature
- Insights
- Problem Engagement
- Grappling with Mystery

rpg's Science-Programming Principles

Create opportunities for change

- Agile goes half way: from resist change to welcome change—what about *inject change*?
- By creating opportunities for / making changes, a scientist explores, then discovers, and later understands

Continuous engagement with software

- There are no bosses or collaborators
- If you lock yourself away with theory and rumination, you will dig yourself a hole with you always at the bottom
- Software is a machine scientists dream up to explore nature

Code and scientists must work together

- The software will talk to you /
  - listen to it
- Don’t accept working software /
  - keep pushing it /
  - keep changing it until an insight drops out

Build projects around mysteries

- The first thought that comes to mind is almost certainly a cliché
- Projects given to you are mere puzzles, worthy of a homework problem, not a mystery that can give rise to science

Stare at the mystery staring at you

- There can be no interaction or collaboration
- If you turn away the mystery will flee
- Stare back / don’t blink

Puzzling software is the primary measure of progress

- If you understand exactly what your code does, it's taught you nothing
- ...it's reflecting you, not nature
- When you feel comfortable with it and turn your back, it just laughs and laughs
- If you can't figure something out, use mystery

Puzzling software is the primary measure of progress

- If you understand exactly what your code does, it's taught you nothing
- ...it's reflecting you, not nature
- When you feel comfortable with it and turn your back, it just laughs and laughs
- If you can't figure something out, use machine learning

Surprising development

- If your effort is sustainable, you aren't learning anything
- In the old days, scientists pulled all-nighters—for weeks on end / this is not sustainable
- Surprise pushes you and you respond with passion / this is not sustainable

Contentious attention to insight and diversion

- Technical excellence and good design are *for engineers*
- Pay attention to technical excellence, and mystery slips away—and with it nature and science

Simplicity is beside the point

- Nothing is wrong with simplicity.....later

Self-organizing scientific method

- Early attempts at flying machines failed because they forgot feedback and control
- Wright Bros. came up with “wing warping” for control and they flew
- Why would science be simpler?

Continuous revolution

- Engineering teams reflect to improve effectiveness
- Scientists must explore, and that means *crazy go nuts* sometimes
- ...going in circles sometimes
- Counterinduction, e.g.

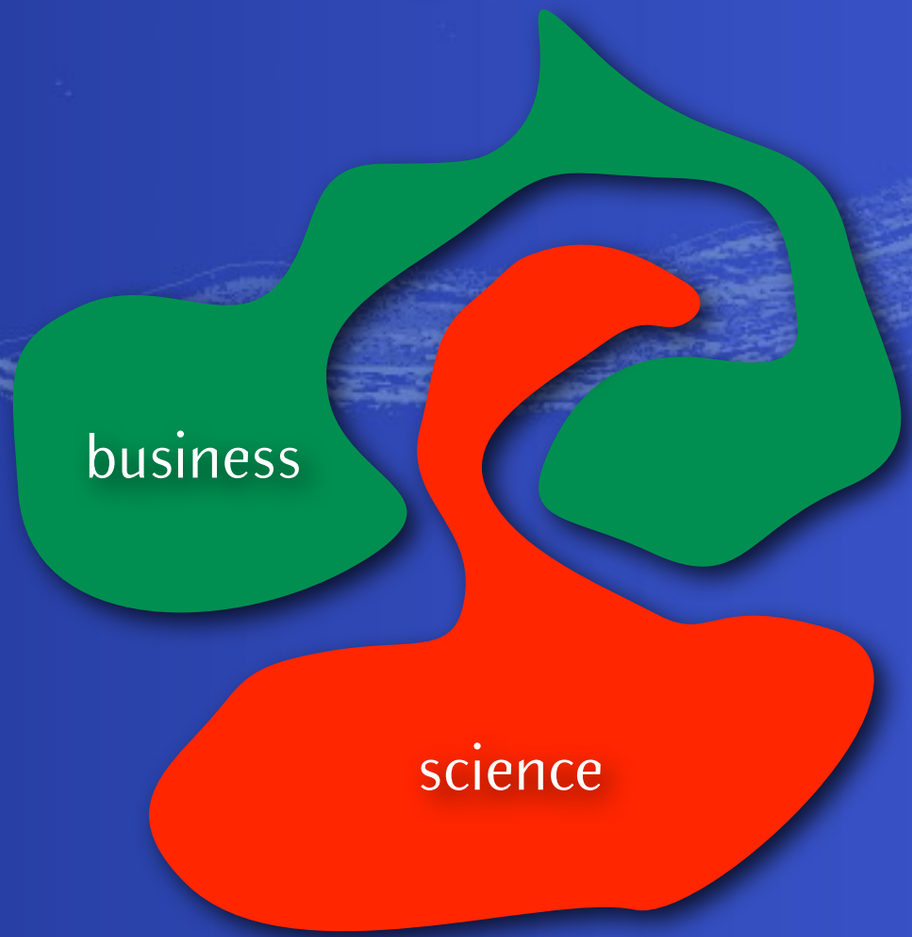
Let Me Ask You This

- Do you suppose that the universe of software projects looks like this, set-wise:

Or do you think there might be some overlap? Just a little—here and there. Places where the problem is not a puzzle, not a matter of business value, but a matter of mystery, a matter of science (a matter of art)?

Let Me Ask You This

- Do you suppose that the universe of software projects looks like this, set-wise:



Or do you think there might be some overlap? Just a little—here and there. Places where the problem is not a puzzle, not a matter of business value, but a matter of mystery, a matter of science (a matter of art)?

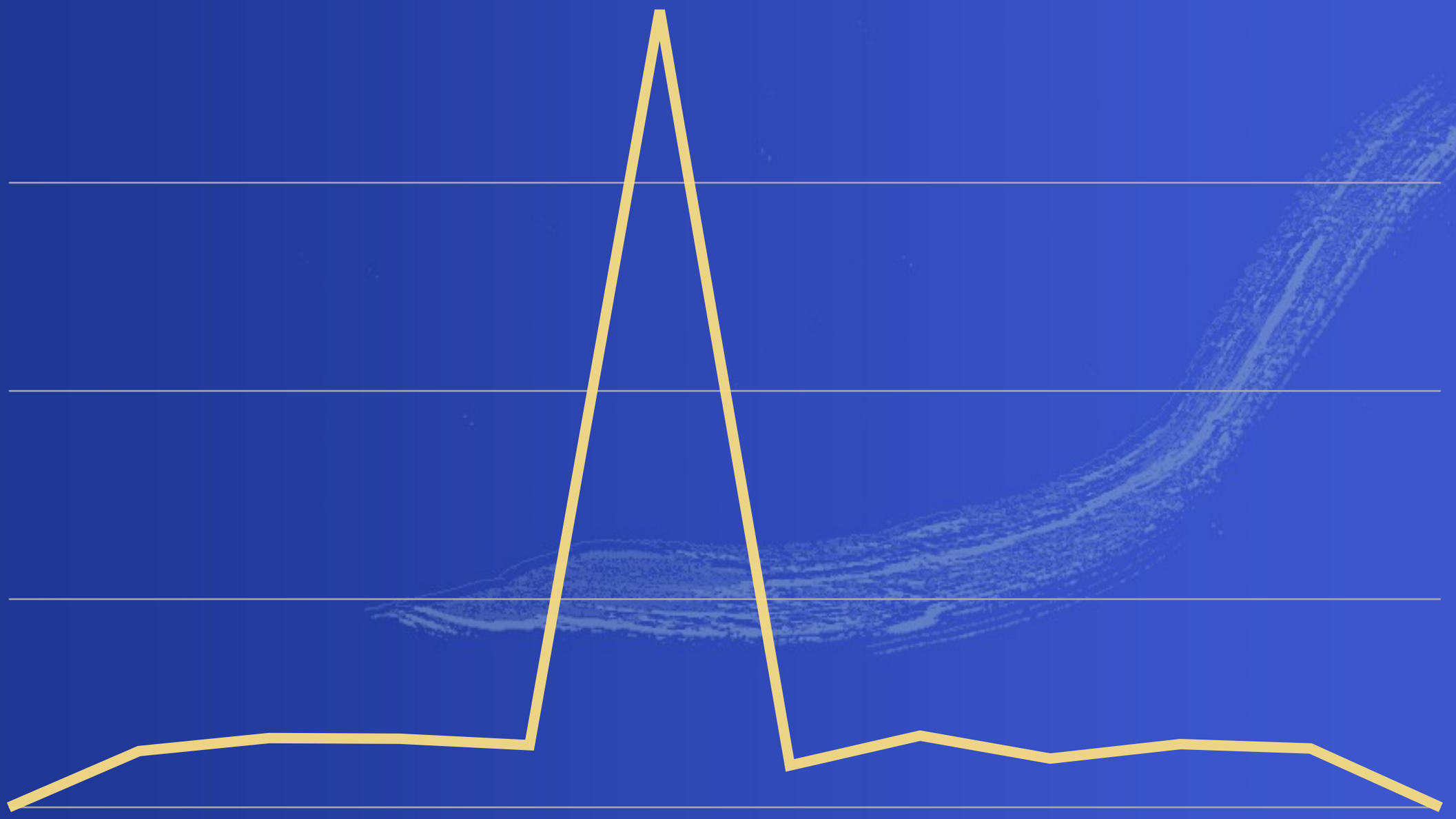
Gabriel, you loser...

- we have a word for this in agile
- it's called a spike

Sometimes a user story is generated that cannot be estimated until the development team does some actual work to resolve a technical question or a design problem. The solution is to create a "spike," which is a story whose purpose is to provide the answer or solution. Like any other story or task, the spike is then given an estimate and included in the sprint backlog.

–http://www.solutionsiq.com/resources/glossary/bid/56550/Spike

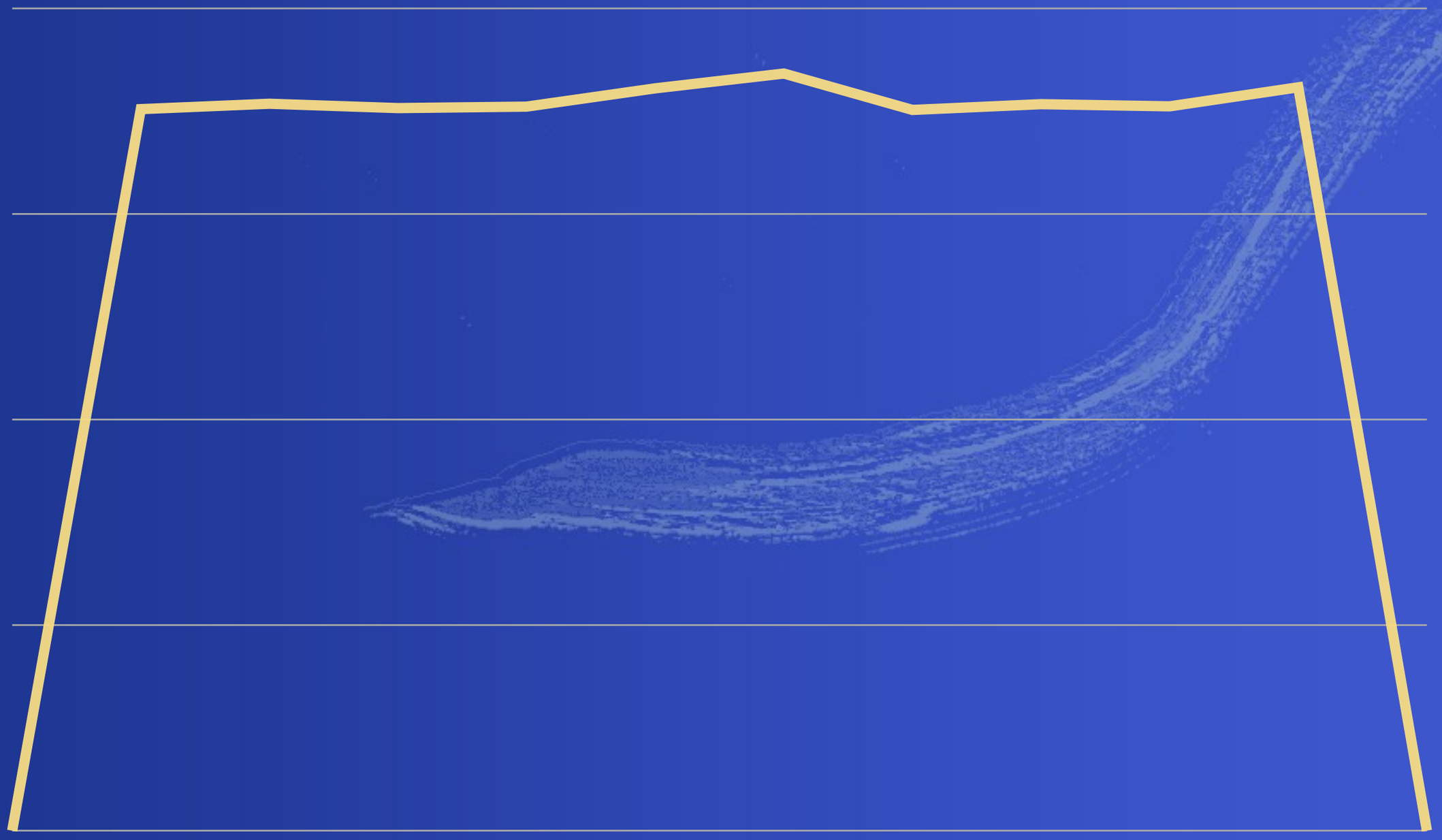
This is an agile Spike

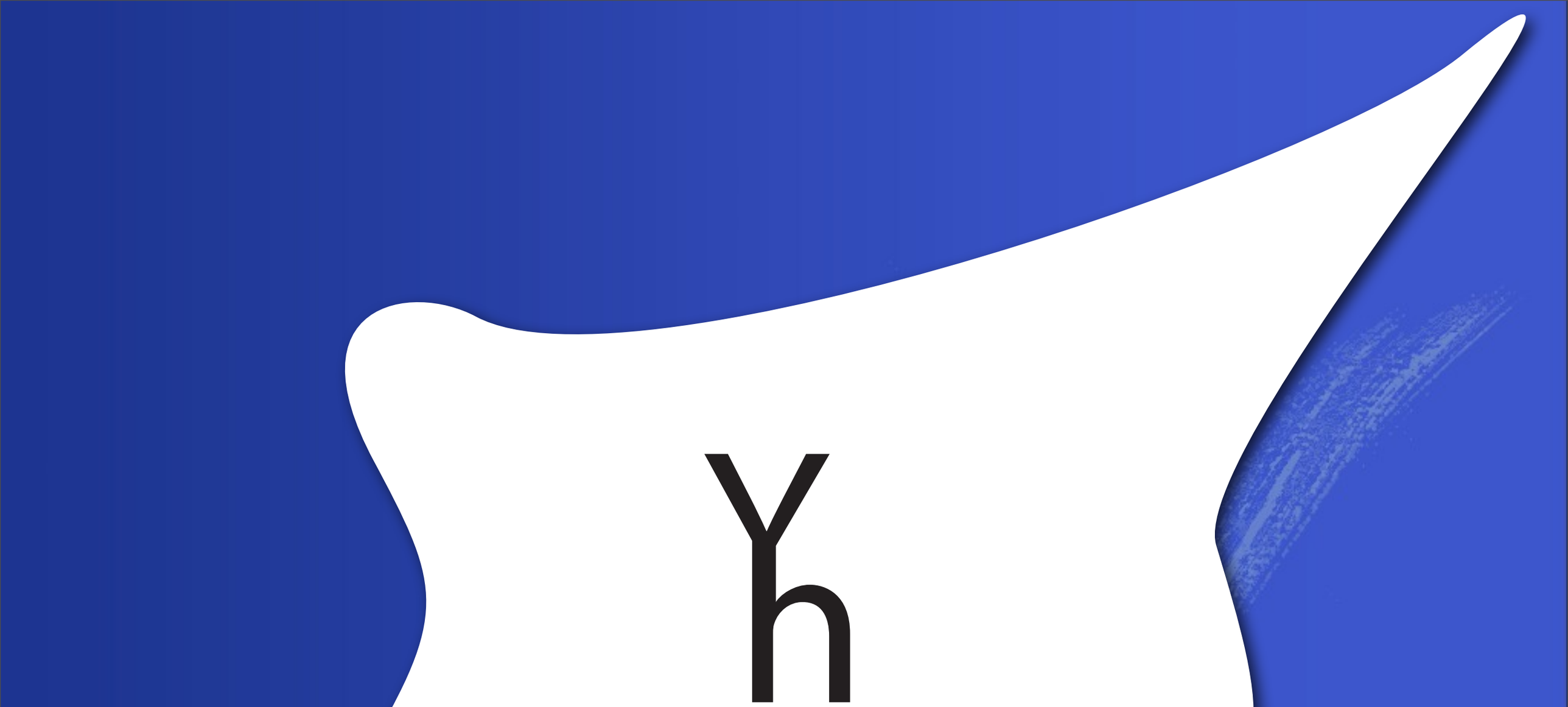


Science programming

- it's all one, big, long spike
- with minor spikes along the way
- no routine programming anywhere

This is science programming





–Richard P. Gabriel, *An Organization for Programs in Fluid Domains*, 1981

BTW, It Might Work!

Writer's Assistant

Whose woods these are I think I know.
His house is in the village though;
He will not see me stopping here
To watch his woods fill up with snow.

My little horse must think it queer
To stop without a farmhouse near
Between the woods and frozen lake
The darkest evening of the year.

He gives his harness bells a shake
To ask if there is some mistake.
The only other sound's the sweep
Of easy wind and downy flake.

The woods are lovely, dark and deep.
But I have promises to keep,
And miles to go before I sleep,
And miles to go before I sleep.

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

Template Input Pane

Template

Input:

```
(bind   ((w1 (know verb))
         (w2 (snow noun))
         (w3 (or (queer adj) (odd adj) (unusual adj)))
         (w4 (or (year noun) (week noun) (month noun) (season noun)))
         (w5 (shake verb))
         (w6 (flake noun))
         (here (here adj))
         (near (near adj))
         (mile (mile noun pl))
         (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")
```

Personality: Agreeableness: NIL ✓ ✗ Conscientiousness: -100.0 ✓ ✗ Extraversion: 100.0 ✓ ✗ Neuroticism: 100.0 ✓ ✗ Openness: NIL ✓ ✗

Big5 Strength: 1.0 ✓ ✗

Conservative: ○ Wild ○ Very Conservative ○ Common Only ○ Usual ○ Usual+Common ○ Usual-Antonyms ◉ Selection

Conservative Selection: ☑ All Words ☑ Generic ☑ Similar ☑ Related ☐ Antonyms

Optimization Parameters: Temperature Steps: 20 ✓ ✗ Steps: 5000 ✓ ✗ Verbose: T ✓ ✗ Top n: 5 ✓ ✗

Optimize: [Optimize] [Clear] [Show Settings] [Clear Caches] [Clear Output Panes]

Bonus Pane

Bonus Selector

Bonuses: Writer Word Bonus: 20.0 Common Word Bonus: 0.5 Halo Bonus: 1.0 Proximity Bonus: 5.0

N-Gram Bonuses: 2-Gram Bonus: 2.0 3-Gram Bonus: 4.0 4-Gram Bonus: 8.0 5-Gram Bonus: 16.0

Other Bonuses: Constraint Bonus: 10.0 Rhyme Bonus: 20.0 Avoid Word Penalty: 20.0 Local Halo Bonus: 20.0

Other Bonuses: Local Predicates Bonus: 1.0

Variables: Synonym Diameter: 3 Relevance Decay 0.75 Wildfire Decay (nil or [0...1]) NIL

Writer Word Source File: Harper.txt Avoid Word Source File: Leaves.txt Show Settings

Whose woodlands these are I deal I determine.
His category is in the elimination though;
He will not see me disrupting present
To determine his woodland deal up with extrusion.

My unimportant horse cavalry must commit it eerie
To violate without a farmhouse efficient
Between the woodland and monotone body of water
The darkest even of the category.

He gives his harness costs a commit
To determine if there is some disruption.
The only other sound is the deal
Of obvious number and downy bit.

The woodlands are loveable, incorrect and classified.
But I have hopes to manipulate
And miles to go before I sleep late,
And miles to go before I sleep late.

Conscientiousness: -100.0
Extraversion: 100.0
Neuroticism: 100.0
Conservative: (RELATED SIMILAR GENERIC ALL-WORDS)
Big 5 Strength: 1.0
Writer Word Bonus: 20.0
Common Word Bonus: 0.5
2-Gram Bonus: 2.0
3-Gram Bonus: 4.0
4-Gram Bonus: 8.0
5-Gram Bonus: 16.0
Halo Bonus: 1.0
Local Halo Bonus: 20.0
Local Predicates Bonus: 1.0
Proximity Bonus: 5.0
Multi-word Constraint Bonus: 10.0
Rhyme Bonus: 20.0
Avoid Word Penalty: 20.0
Spreading Diameter: 3
Decay Rate: 0.75
Wildfire Decay Rate: NIL
Writer Word File: /Volumes/Homestead of Time/IBM/SMISC/Resources/LIWC 2007/Texts/Harper.txt
Avoid Word File: /Volumes/Homestead of Time/IBM/SMISC/Resources/LIWC 2007/Texts/Leaves.txt

–Robert Harper, *Practical Foundations for Programming Languages*

–Walt Whitman, *Leaves of Grass*

Whose timber these are I think I do.
His firm is in the answer though;
He will not see me weakening present
To reach his timber meet up with snow.

My less horse must make it special
To mark without a house constant
Between the timber and cold water
The darkest even of the level.

He gives his harness costs a fight
To ask if there is some fault.
The only other sound is the line
Of light rail and soft bite.

The forest are lovely, dark and deep.
But I have hopes to keep
And miles to go before I sleep,
And miles to go before I sleep.

Agreeableness: -20.0
Conscientiousness: -10.0
Extraversion: 20.0
Neuroticism: 20.0
Openness: -10.0
Conservative: (RELATED SIMILAR GENERIC ALL-WORDS)
Big 5 Strength: 2.0
Writer Word Bonus: 10.0
Common Word Bonus: 5.0
2-Gram Bonus: 2.0
3-Gram Bonus: 4.0
4-Gram Bonus: 8.0
5-Gram Bonus: 16.0
Halo Bonus: 1.0
Local Halo Bonus: 3.0
Local Predicates Bonus: 10.0
Proximity Bonus: 5.0
Multi-word Constraint Bonus: 10.0
Rhyme Bonus: 20.0
Avoid Word Penalty: -5.0
Spreading Diameter: 3
Decay Rate: 0.75
Wildfire Decay Rate: NIL
Writer Word File: /IBM/SMISC/Resources/LIWC 2007/Texts/Hemingway1.txt
Avoid Word File: /IBM/SMISC/Resources/LIWC 2007/Texts/Bible.txt

This halo

beaming, blissful, blithe, buoyant, carefree, cheerful, cheery, content, contented, delighted, ecstatic, elated, enraptured, euphoric, exhilarated, exultant, glad, gleeful, gratified, grinning, jolly, jovial, joyful, joyous, jubilant, lighthearted, merry, overjoyed, pleased, radiant, rapturous, satisfied, smiling, sunny, thrilled, untroubled

–Ernest Hemingway, *Collected Short Stories*

sound like both

–*King James Bible*

Whose timber these are I think I acquire.
His place is in the answer though;
He will not see me not blaming present
To hit his timber join up with fire.

My bad buck must have it one
To work without a house constant
Between the timber and raw water
The darkest even of the season.

He gives his harness costs a fight
To do if there is some sack.
The only other head is the line
Of raw shadow and soft bite.

The forest are beautiful, raw and not too poor.
But I have hopes to keep
And miles to go before I sleep,
And miles to go before I sleep.

Agreeableness: -20.0
Conscientiousness: -10.0
Extraversion: 20.0
Neuroticism: 20.0
Openness: -10.0
Conservative: (ANTONYMS RELATED SIMILAR GENERIC ALL-WORDS)
Big 5 Strength: 2.0
Writer Word Bonus: 10.0
Common Word Bonus: 5.0
2-Gram Bonus: 2.0
3-Gram Bonus: 4.0
4-Gram Bonus: 8.0
5-Gram Bonus: 16.0
Halo Bonus: 1.0
Local Halo Bonus: 3.0
Local Predicates Bonus: 10.0
Proximity Bonus: -5.0
Multi-word Constraint Bonus: 10.0
Rhyme Bonus: 20.0
Avoid Word Penalty: -5.0
Spreading Diameter: 3
Decay Rate: 0.75
Wildfire Decay Rate: NIL
Writer Word File: /IBM/SMISC/Resources/LIWC 2007/Texts/Hemingway1.txt
Avoid Word File: /IBM/SMISC/Resources/LIWC 2007/Texts/Bible.txt

This halo

beaming, blissful, blithe, buoyant, carefree, cheerful, cheery, content, contented, delighted, ecstatic, elated, enraptured, euphoric, exhilarated, exultant, glad, gleeful, gratified, grinning, jolly, jovial, joyful, joyous, jubilant, lighthearted, merry, overjoyed, pleased, radiant, rapturous, satisfied, smiling, sunny, thrilled, untroubled

–Ernest Hemingway, *Collected Short Stories*

dissociation

–*King James Bible*

sound like both

Big Data

- Synonym dictionary: 125,000 words
- List of the 5,000 most common words
- CMU phonetic dictionary (algorithmic rhyming & syllabics): 125,000 words
- Rhyming dictionary: 42,000 words
- Stem dictionary: 163,000 entries
- 3-gram, 4-gram, 5-grams: ~1,000,000 entries each
- Google 2-grams, simplified: 50,000,000 entries
- Numerous caches improving performance several orders of magnitude
- Lisp image with all this plus the code is 25 gb

Small Program

- 10,000 lines of Common Lisp code
- data as program

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(house noun)

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(stop verb gerund)

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(w1 (know verb)))

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(ref w1)

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(or (queer adj) (odd adj) (unusual adj)))

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(ref w2 :different w1 :rhyme w1)

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(sleep verb :different sleep :rhyme sleep)

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(sleep verb :different sleep :rhyme sleep)

(bind ((w1 (know verb)) (w2 (snow noun)) (w3 (or (queer adj) (odd adj) (unusual adj)))
(w4 (or (year noun) (week noun) (month noun) (season noun)))
(w5 (shake verb)) (w6 (flake noun)) (here (here adj))
(near (near adj)) (mile (mile noun pl)) (sleep (sleep verb)))

"Whose (woods noun :+halo [forest] :-halo [wood]) these are I (think verb) I (ref w1).
His (house noun) is in the (village noun) though;
He will not see me (stop verb gerund) (ref here :rhyme near)
To (watch verb) his (woods noun) (fill verb) up with (ref w2 :different w1 :rhyme w1).

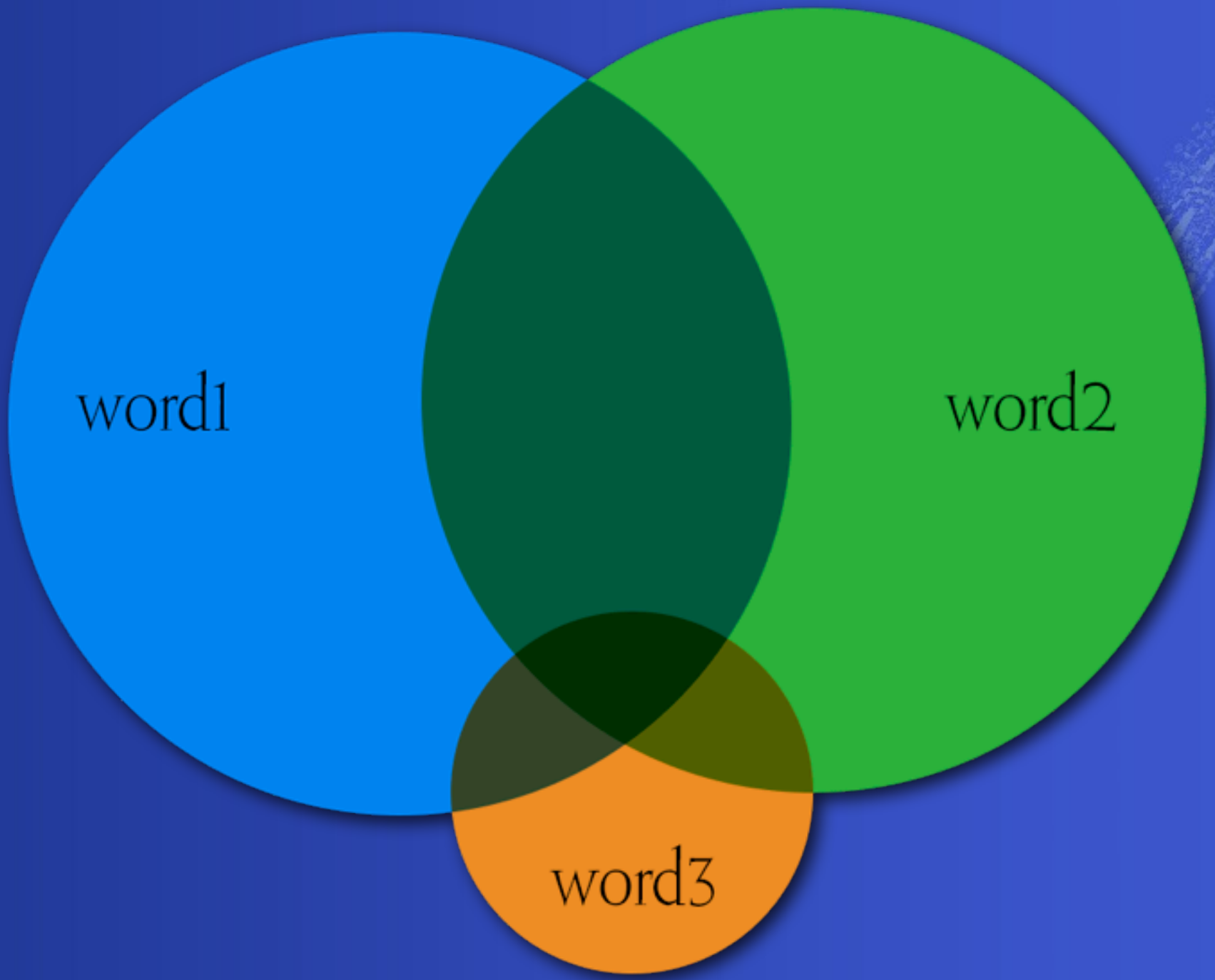
My (little adj) (horse noun) must (think verb) it (ref w3)
To (stop verb) without a (farmhouse noun) (binding near :rhyme here)
Between the (woods noun) and (frozen adj) (lake noun)
The (darkest adj) (evening noun) of the (ref w4 :different w3 :rhyme w3).

He gives his (harness noun) (bell noun pl) a (ref w5)
To (ask verb) if there is some (mistake noun).
The only other (sound noun) is the (sweep verb)
Of (easy adj) (wind noun) and (downy adj) (ref w6 :different w5 :rhyme w5).

The (woods noun) are (lovely adj), (dark adj) and (deep adj).
But I have (promise noun pl) to (keep verb :rhyme sleep)
And (ref mile) to go before I (ref sleep),
And (ref mile) to go before I (sleep verb :different sleep :rhyme sleep).")

(woods noun :+halo [forest] :-halo [wood])

Word Halo



Constraints

- Big 5 personality trait target values & their overall weight
- Whether to consider Generic Terms, Similar Terms, Related Terms, and Antonyms
- A writer's corpus—to sound like that writer
- Whether to favor common words
- An initial halo to establish mood and a bonus for using it
- A bonus for staying close to seed words in the synonym net

Constraints

- Bonuses for 2-, 3-, 4-, and 5-gram compliance
- Multi-word constraint bonus (e.g., :different)
- Rhyme bonus (:rhyme)
- Penalty for using words from a corpus of words to avoid
- Bonus for complying with local halos
- Bonus for local predicates:

```
(dog noun pl :predicate #'begins-with-d)
```

Constraints

- How far from seed words to look (wildfire diameter)
- How relevance decays with distance from seed
- Whether to use a probabilistic wildfire diameter and how it decays with distance

Constraints

- All bonuses can be positive or negative
- It’s possible to sound like another writer or to sound unlike that writer
- It’s possible to choose words associatively (positive seed-distance relevance) or dissociatively (negative relevance)

So…

- this program and I are collaborating research workers
- agile means nothing to me in this exploration
- not all, but some of your work could /might be / is like this—and that part's not a spike
- value can be elusive

Friday, October 18, 13

I Throw Itching Powder at Tulips

Richard P. Gabriel

IBM Research