



Enhancing document representations using analysis of content difficulty

Models, Applications, and Insights

Kevyn Collins-Thompson

Associate Professor
University of Michigan

School of Information &
College of Engineering, EECS

© 2001-21 K. Collins-Thompson

Thanks to my valued collaborators:

<u>Rohail Syed</u>	<i>Univ. of Michigan</i>
<u>Karthik Raman</u>	<i>Cornell / Google Research</i>
Gwen Frishkoff	<i>University of Oregon</i>
<u>Elliot Schumacher</u>	<i>Carnegie Mellon / JHU</i>
Jamie Callan	<i>Carnegie Mellon</i>
Maxine Eskenazi	<i>Carnegie Mellon</i>
Sebastian de la Chica	<i>Microsoft</i>
Paul Bennett	<i>Microsoft Research</i>
Susan Dumais	<i>Microsoft Research</i>
Ryen White	<i>Microsoft Research</i>
Michael Heilman	<i>Educational Testing Service</i>
<u>Paul Kidwell</u>	<i>Purdue / LLNL</i>
Guy Lebanon	<i>Georgia Tech / Google</i>
David Sontag	<i>MIT</i>

And the KDD Document Intelligence Workshop organizers

Documents on a topic can occur at a wide range of reading difficulty levels

Grasshopper Habitat and Grasshopper Diet

Grasshoppers live in fields, meadows and just about anywhere they can find generous amounts of food to eat. A grasshopper has a hard shell and a full grown grasshopper is about one and a half inches, being so small you would not think they would eat much - but you would be so wrong - they eat lots and lots - an average grasshopper can eat 16 time its own weight.

The grasshoppers favourite foods are grasses, leaves and cereal crops. One particular grasshopper - the Shorthorn grasshopper only eats plants, but it can go berserk and eat every plant in sight - makes you wonder where they put it all.



Grasshopper Behaviour

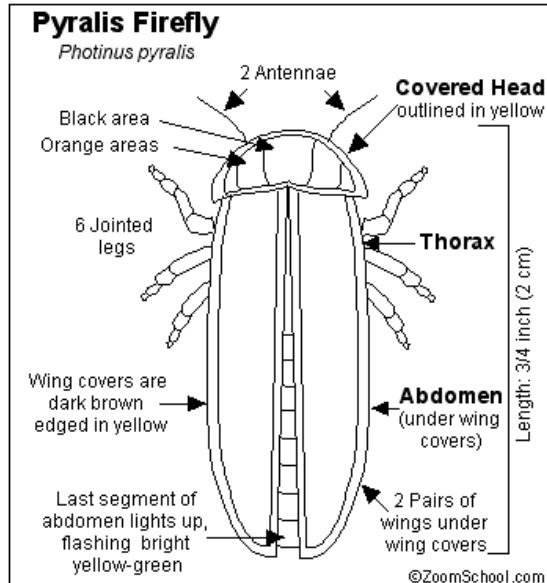
Query [insect diet]: Lower difficulty

Medium difficulty [insect diet]

[Insect Printouts](#)

Firefly or Lightning Bug *Photinus pyralis*

[More Printouts](#)



The Pyralis firefly (also known as the lightning bug) is a common firefly in North America. This partly nocturnal, luminescent beetle is the most common firefly in the USA.

The Firefly's Glow: At night, the very end (the last abdominal segment) of the firefly glows a bright yellow-green color. The firefly can control this glowing effect. The brightness of a single firefly is 1/40 of a candle. Fireflies use their glow to attract other fireflies. Males flash about every five seconds; females flash about every two seconds. This firefly is harvested by the biochemical industry for the organic compounds luciferin (which is the chemical the firefly uses for its bioluminescence).

Anatomy: This flying insect is about 0.75 inch (2 cm) long. It is mostly black, with two red spots on the head cover; the wing covers and head covers are lined in yellow. Like all insects, it has a hard exoskeleton, six jointed legs, two antennae, compound eyes, and a body divided into three parts (the head, thorax, and abdomen).

Diet: Both the adults and the larvae are **carnivores** (meat-eaters). They eat other insects (including other fireflies), insect larvae, and snails.

Higher difficulty [insect diet]

R&D

INSECT REARING

RESEARCH and DEVELOPMENT

at NCSU

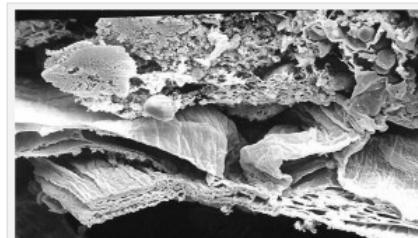
REARING RESEARCH ON DIET DEVELOPMENT AND ESTABLISHING NEW AND IMPROVED REARING SYSTEMS

- Development of artificial diets and rearing systems for many species of insects has been our specialty.
- We use a variety of techniques to help us develop appropriate diets, including analysis of the natural foods, feeding biology of the target insects, bioassays, biological/biochemical testing, and analysis of internal biology.

Some of our recent and current projects are below:



Here are cactus moth larvae (*Cactoblastis cactorum*: Lepidoptera: Pyralidae) developing on one of our newest artificial diets developed for USDA, APHIS and Florida Department of Plant Industry.



The above electron microscope image shows the peritrophic matrix (PM) of a tobacco budworm fed plant parts from its natural diet: note the multiple layers formed in response to a natural food.

Users also exhibit a wide range of proficiency and expertise

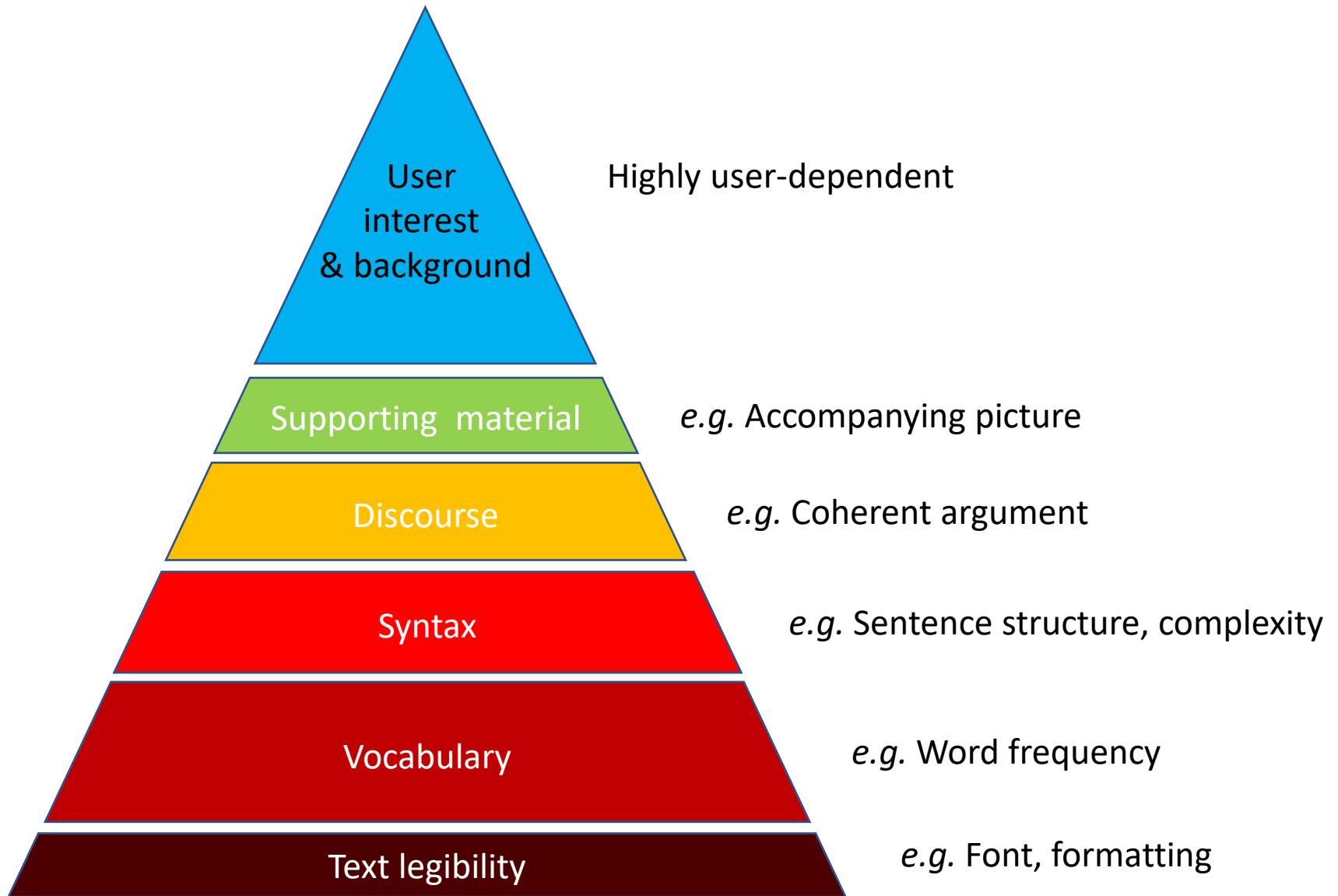
- Students at different grade levels
- Non-native speakers
- General population
 - Overall language proficiency or literacy
 - Familiarity or expertise in specific topic areas

Long-term goal: Optimally connecting users & information for personalized learning and discovery

1. How can we create richer representations of user and content for supporting learning (at Web scale)?
2. How can we integrate models of human learning and cognition into search and recommender algorithm objectives, features, and evaluation?
3. How can search engine algorithms support important educational goals like robust long-term retention, or increase in curiosity...not just short-term learning?

This talk: Analyzing content difficulty enhances document representations for understanding and supporting readers.

What makes text difficult to read and understand?



Traditional readability measures don't work for Web and other non-traditional content

- Flesch-Kincaid (Microsoft Word)

$$RG_{FK} = 0.39 \cdot [Words / Sentence] + 11.8 \cdot [Syllables / Word] - 15.59$$

- Problems include:
 - They assume the content has well-formed sentences
 - They are sensitive to noise
 - Input must be at least 100 words long
- Web and other content is often short, noisy, less structured
 - Page body, titles, snippets, queries, captions, ...
 - Health questionnaires, surveys
- Billions of pages → computational constraints on approaches
- We focus on generative vocabulary-based prediction models that learn fine-grained models of word usage from labeled texts

Model 1: Vocabulary-based generative models: smoothed unigram language models

[**Collins-Thompson & Callan: HLT 2004**]

1. Model each grade G_i as a word histogram θ_i .
2. Smooth θ_i by combining evidence from nearby grades.

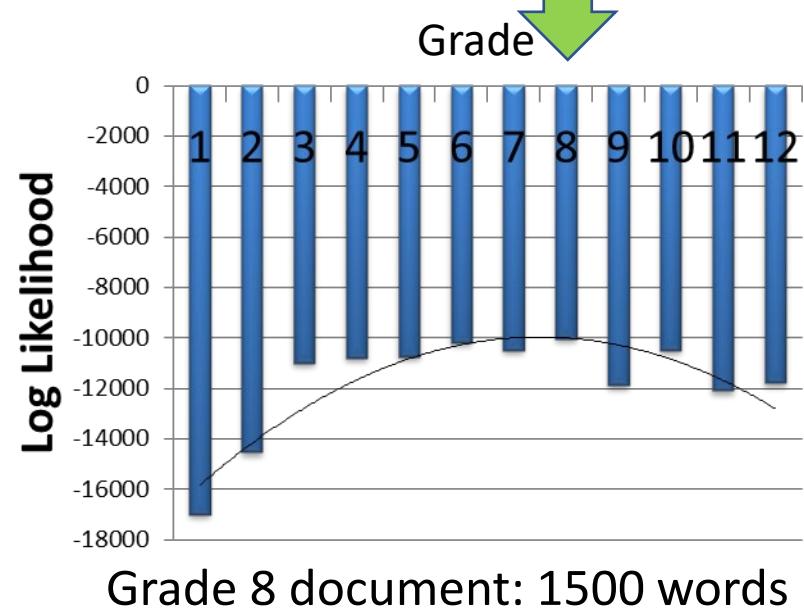
$$\hat{\theta}_i(w) = \sum_{j=1 \dots |G|} \phi_h(i, j) \hat{P}(w|G_j)$$

3. Compute the likelihood of the text in each grade model θ_i .

$$\log P(G_i|T) = \sum_{w \in T} C(w) \log \hat{\theta}_i(w) + \log Q$$

4. Prediction: Select the **most likely** grade.

Type	Grade 1 $P(w G_1)$	Grade 5 $P(w G_5)$	Grade 12 $P(w G_{12})$
the	0.080	0.090	0.100
a	0.060	0.050	0.060
red	0.020	0.005	0.0007
ball	0.010	0.0001	0.00005
was	0.010	0.010	0.200
perimeter	0.003	0.040	0.004
optimized	0.00001	0.0001	0.010



Which words are most ‘distinctive’ of each grade in these language models?

Grade 1

grownup	2.485
ram	2.425
planes	2.411
pig	2.356
jimmy	2.324
toad	2.237
shelf	2.192
cover	2.184
spot	2.174
fed	2.164

Grade 4

desert	1.787
crew	1.765
habitat	1.763
butterflies	1.758
rough	1.707
slept	1.659
bowling	1.643
ribs	1.610
grows	1.606
entrance	1.604

Grade 8

acidic	1.425
soda	1.425
acid	1.408
typical	1.379
angle	1.362
press	1.318
radio	1.284
flash	1.231
levels	1.229
pain	1.220

Grade 12

essay	2.441
literary	2.383
technology	2.363
analysis	2.301
fuels	2.296
senior	2.292
analyze	2.279
management	2.269
issues	2.248
tested	2.226

*These values are computed using a Fisher information-type statistic

Model 2: Estimate word acquisition events

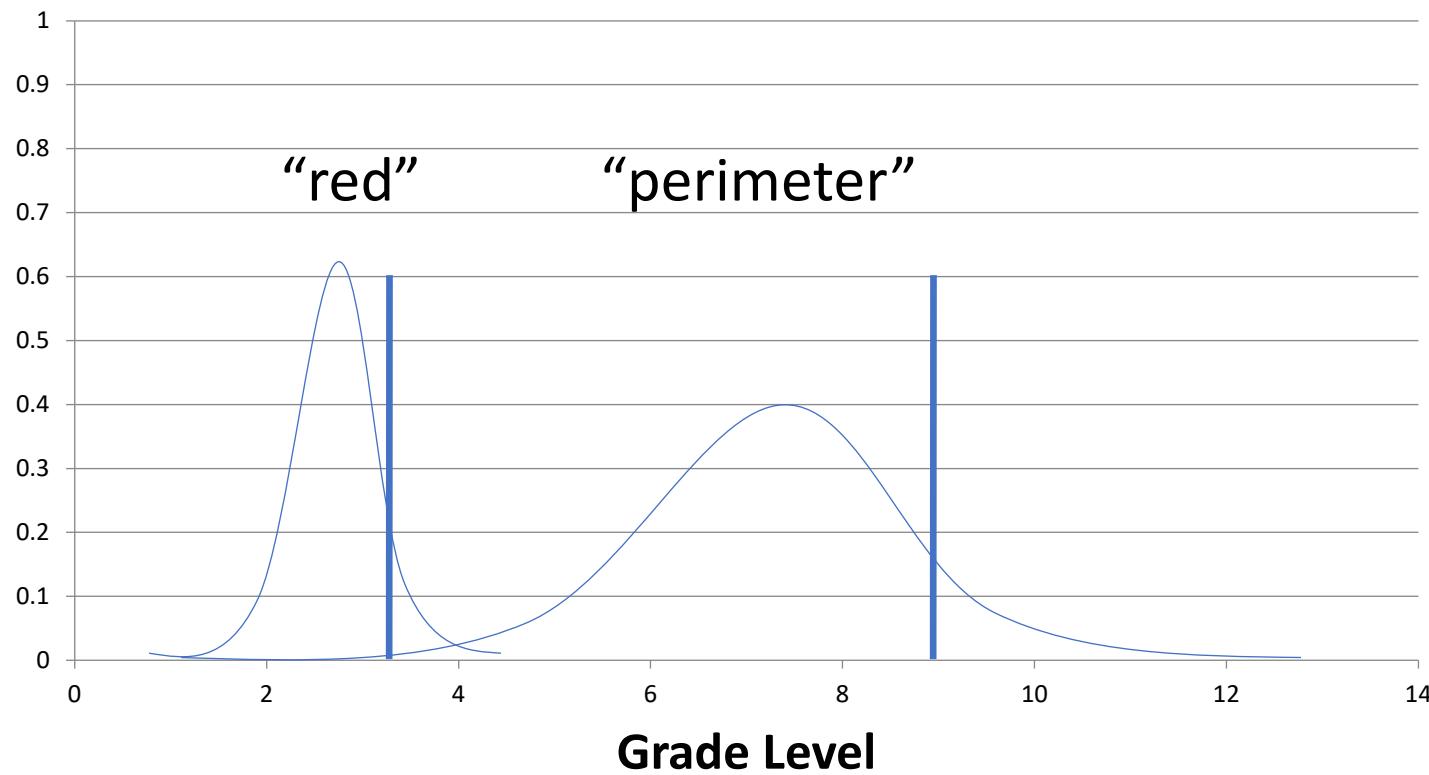
[Kidwell, Lebanon, Collins-Thompson EMNLP 2009, J. Am. Stats 2011]

- Inspiration: Dale-Chall list of 3000 words familiar to 80% of 4th -graders.
- Documents can contain high-level words but still be low-level, e.g. teaching new concepts
- (r, s) readability:
 - r : familiarity threshold for any word
A word w is familiar at a grade if known by at least r percent of population at that grade
 - s : coverage requirement for documents
A document d is readable at level t if s percent of the words in d are familiar at grade t .
- When does someone learn to read word w ?
 - Average acquisition age μ_w with standard deviation σ_w (Gaussian)
 - Fit Gaussian (μ_w, σ_w) parameters for each word w
 - Learn all parameters by maximum likelihood from labeled documents
- (r, s) parameters allow tuning the model for different scenarios

The r parameter controls the familiarity threshold for words

Level quantile for word w : $q_w(r)$

$$q_{\text{RED}}(0.80) = 3.5 \quad q_{\text{PERIMETER}}(0.80) = 8.2$$

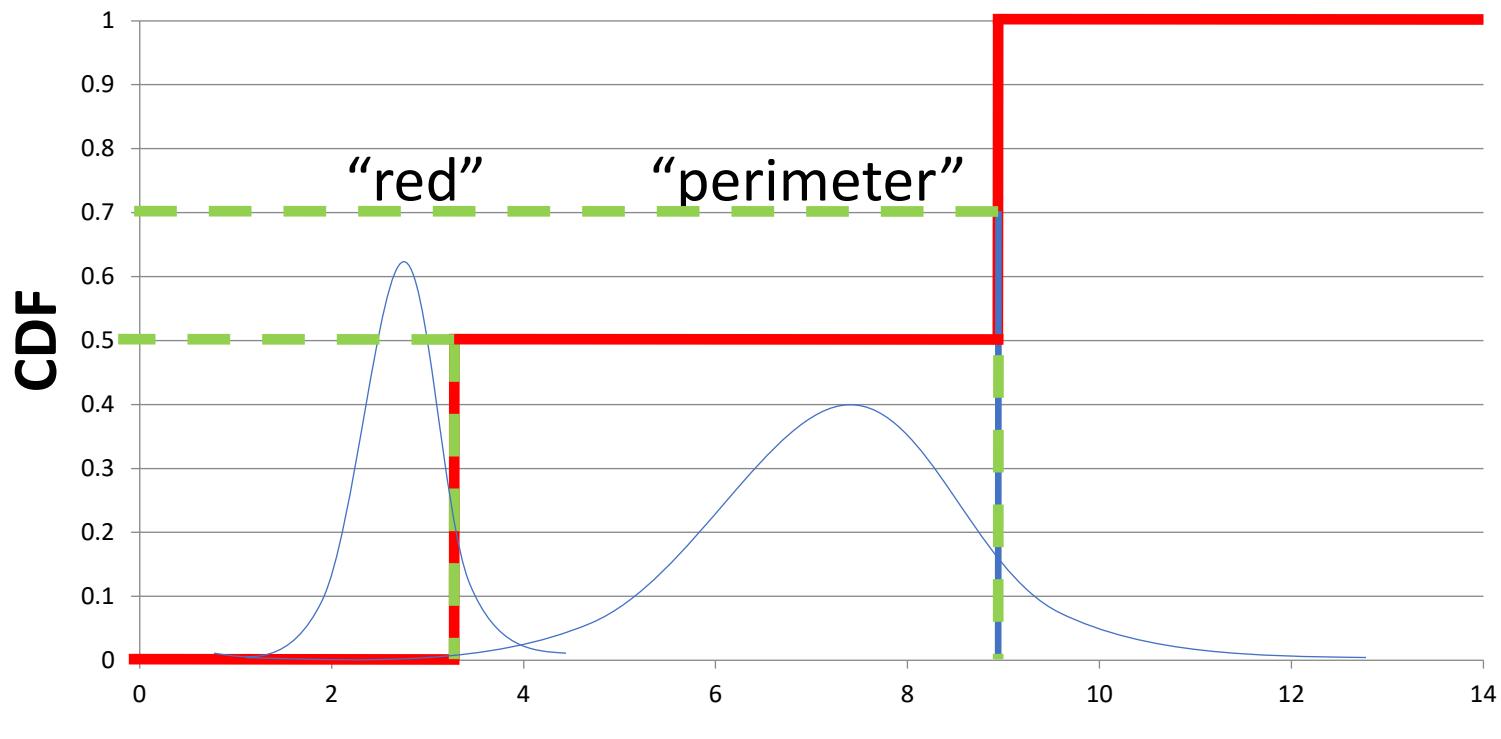


The s parameter controls required document coverage

Suppose: $p(\text{"red"} \mid d) = p(\text{"perimeter"} \mid d) = 0.5$

Predicted grade with $s = 0.70$: 8.8

Predicted grade with $s = 0.50$: 3.5

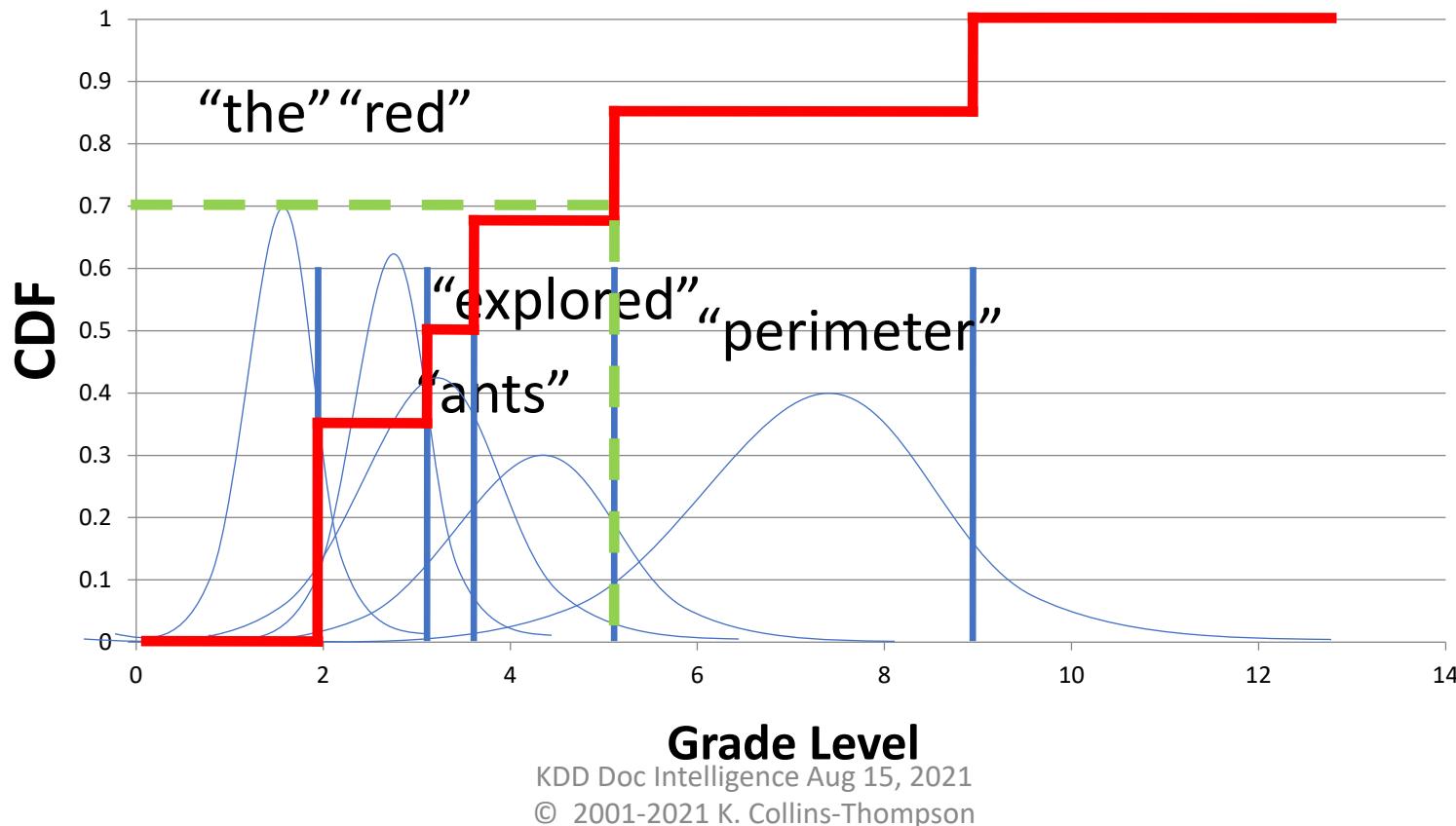


Grade Level

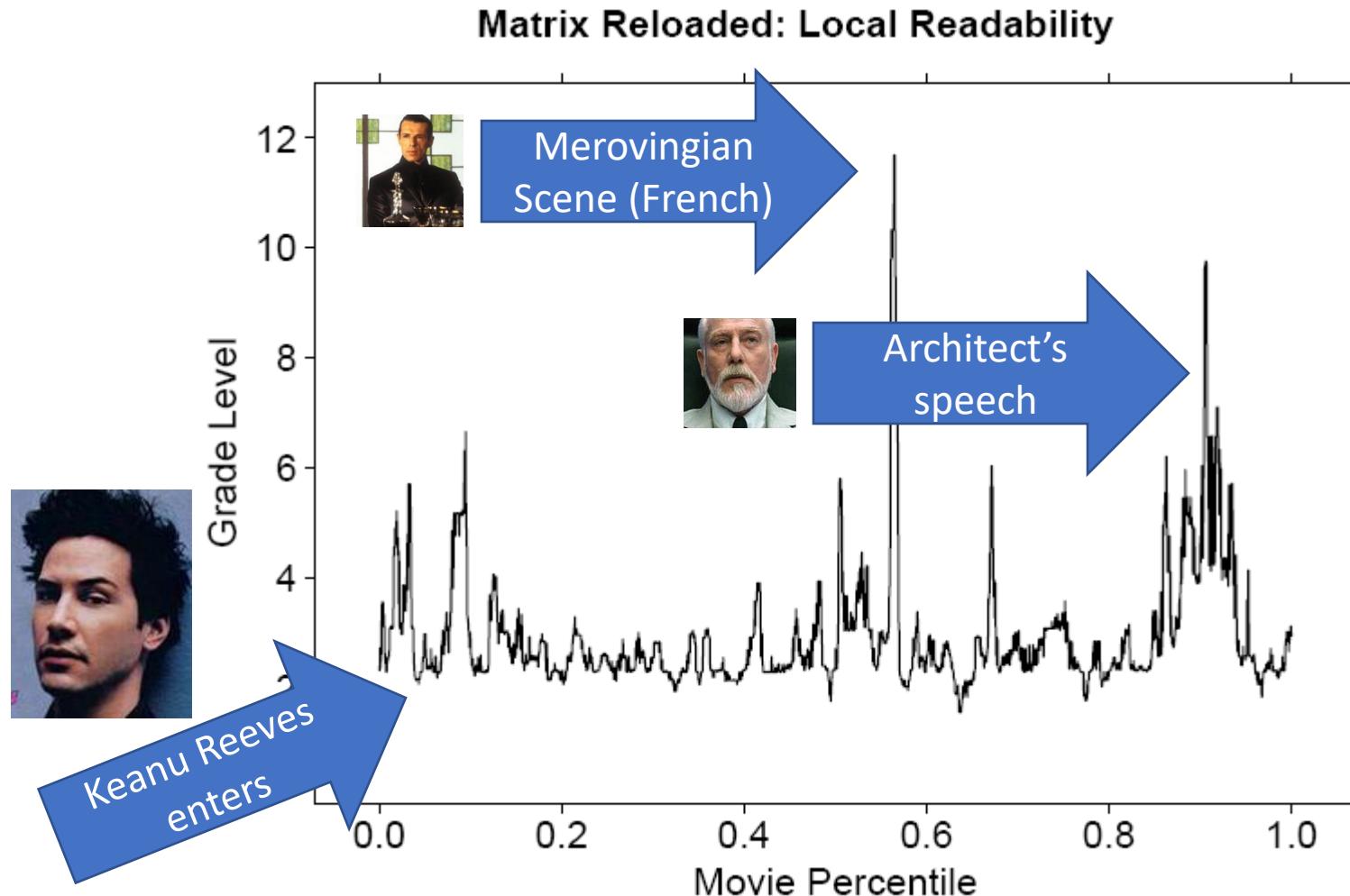
Multiple-word example

“The red ants explored the perimeter.”

Predicted grade with $s = 0.70$: 5.3



Local readability for documents with varying difficulty Movie dialogue in “The Matrix: Reloaded”

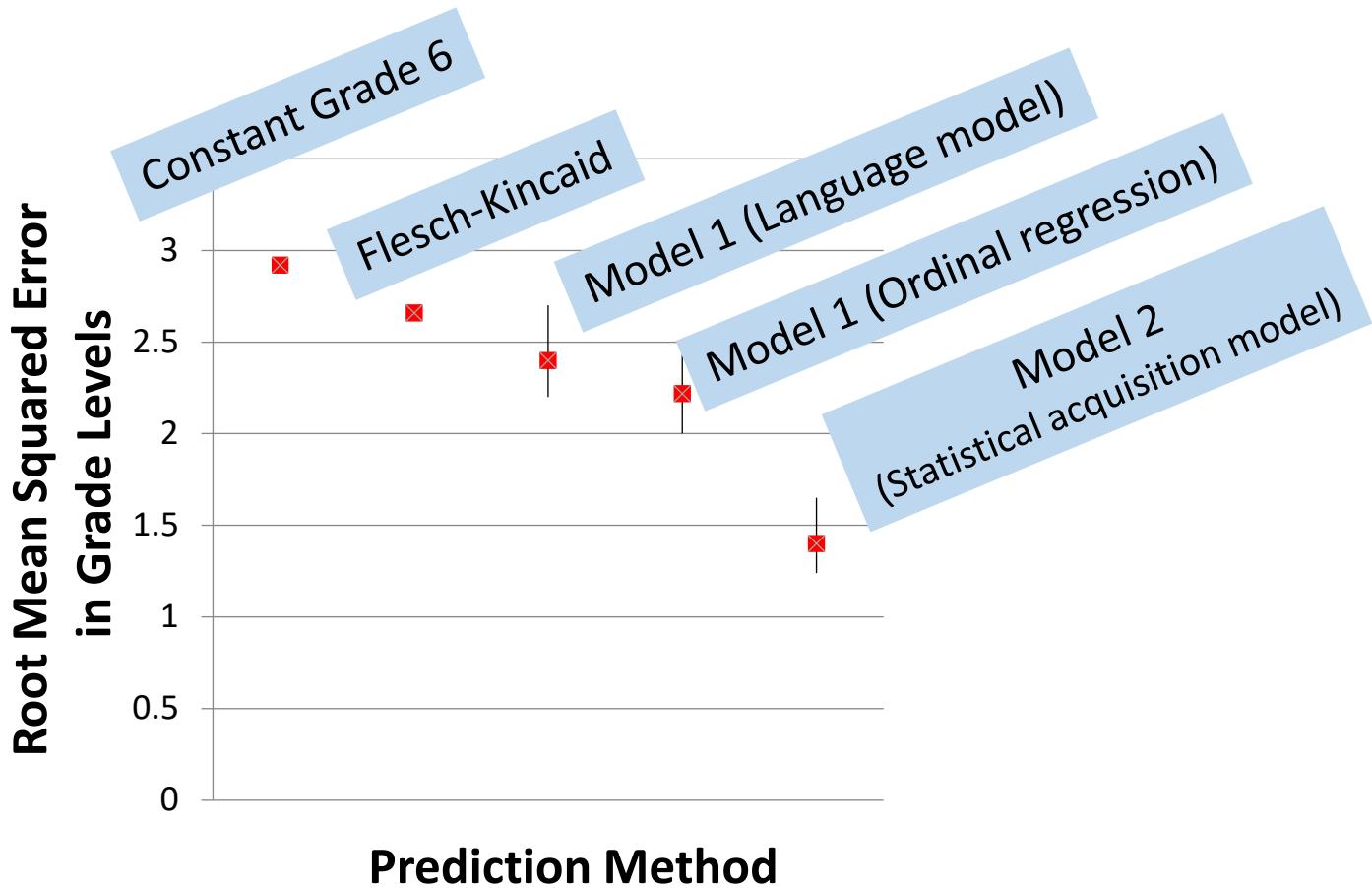


[Kidwell, Lebanon, Collins-Thompson. J. Am. Stats. 2011]

KDD Doc Intelligence Aug 15, 2021

© 2001-2021 K. Collins-Thompson

More detailed vocabulary models improve prediction accuracy for Web content (lower is better)



Selected extensions for readability prediction methods I've explored

- First- vs second-language learners

[M. Heilman, **K. Collins-Thompson**, J. Callan and M. Eskenazi. HLT 2007]

- Rich feature spaces: vocabulary, syntax

[M. Heilman, **K. Collins-Thompson** and M. Eskenazi. ACL BEA workshop 2008]

- Crowdsourcing reliable difficulty labels

[X. Chen, P.N. Bennett, **K. Collins-Thompson**, E. Horvitz. WSDM 2013]

- Single-sentence readability (relative difficulty)

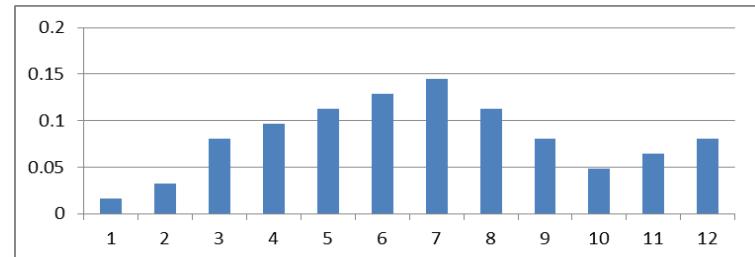
[E. Schumacher, M. Eskenazi, G. Frishkoff, **K. Collins-Thompson**. EMNLP 2016]

See my computational readability survey linked on my UMichigan homepage
International Journal of Applied Linguistics 165:2

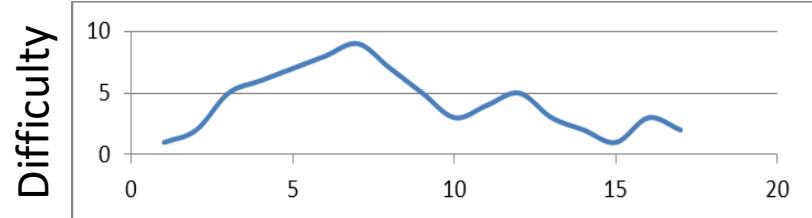
Reading difficulty provides a rich new representation of documents, sites, and users

Level 1: Documents

- Distribution over levels
- Can find key ‘words to learn’
- View change over time/ position



Health article: *Bronchitis, efficacy ...*



Level 2: Web sites

- e.g. Reading level mean and variance across site pages

Level 3: User profiles

- Reading level via user clicks, visits to documents and sites

What happens when you can label billions of Web pages with reading difficulty metadata?

Topic drift can occur when the specified reading level changes .
Example: [quantum theory]

[Quantum mechanics - Wikipedia, the free encyclopedia](#)

[History](#) · [Mathematical formulations](#) · [Mathematically ...](#) · [Interactions with ...](#)

Quantum mechanics (QM - also known as quantum physics, or quantum theory) is a branch of physics dealing with physical phenomena where the action is on the order ...
en.wikipedia.org/wiki/Quantum_mechanics

[quantum theory: Definition from Answers.com](#)

[quantum theory](#) n. A theory in physics based on the principle that matter and energy have the properties of both particles and waves, created to explain
www.answers.com/topic/quantum-theory

[Quantum theory - Wikipedia, the free encyclopedia](#)

Quantum theory may mean: In science: Quantum mechanics: a subset of quantum physics explaining the physical behaviours at atomic and sub-atomic levels Old quantum ...
en.wikipedia.org/wiki/Quantum_theory

[Quantum Theory - thebigview.com - Pondering the Big Questions](#)

Discovering the fundamental structure of matter. Quantum theory evolved as a new branch of theoretical physics during the first few decades of the 20th century in an ...
www.thebigview.com/spacetime/quantumtheory.html

Top 4 results

[quantum theory] + lower difficulty

[Quantum Theory - PS3 - IGN - Sony PlayStation 3 ...](#)

PlayStation 3 · 29 photos · Walkthroughs · Cheats

Sep 28, 2010 · **Quantum Theory** is a game whose design is dated

despite being a week old. It's a game that feels like it didn't ...

ps3.ign.com/objects/142/14288075.html



2.5/10

score

[Quantum Theory : Mix That Drink](#)

I wonder where the **Quantum Theory** cocktail got its name. There's nothing incomprehensible about this cocktail, and it's not as mind-blowing as, say, the Zombie. mixthatdrink.com/quantum-theory

[Quantum Theory Cheats, Codes, and Secrets for PlayStation 3 - GameFAQs](#)

For **Quantum Theory** on the PlayStation 3, GameFAQs has 51 cheat codes and secrets.

www.gamefaqs.com/ps3/954470-quantum-theory/cheats

[Quantum Theory Cheats - Playstation 3 - ActionTrip -- What we lack ...](#)

This page offers the most up-to-date **Quantum Theory** Playstation 3 cheats, codes, and hints. Besides our impressive collection of **Quantum Theory** and other cheats, ...

www.actiontrip.com/cheats/ps3/quantum-theory.phtml

Top 4 results

[quantum theory] + lower difficulty + science topic constraint

Quantum Theory

Quantum theory as a science is officially dead and has been replaced by multiple facets that include such things as **quantum mechanics**. These multi-faceted points ...
www.quantumtheory.org

Does Quantum Theory Explain Consciousness? : Discovery News

Just because consciousness is a mystery and **quantum theory** is mysterious, it doesn't mean they're connected.
news.discovery.com/space/does-quantum-theory-explain-consciousness...

Quantum Theory | PlanetSEED

The Expanding Universe **Quantum Theory** Einstein's Big Mistake? Another big problem goes right back to the way Einstein guessed his equations in 1917.
<https://www.planetseed.com/.../the-expanding-universe/Quantum-Theory>

Einstein's Intuition : Quantum Space Theory

Einstein's Intuition : **Quantum Space Theory**: ... Questions and answers: I'd like to dedicate this page to questions that anyone out there might have regarding

Top 4 results

[cinderella] + higher difficulty

[**Cinderella : Cinderella**](#)

Cinderella is a Java based interactive geometry tool. The only available tool that gives correct solutions to typical geometrical problems.

www.cinderella.de

[**Cinderella Software**](#)

If you only need to browse and/or print SDL files, then download our free viewing tool. Cinderella SDL is a visual modeling tool for developing embedded software ...

www.cinderella.dk/index.htm

[**Cinderella - School of Ballroom Dancing**](#)

About Us: Home | Contact Us : Welcome to the **Cinderella** School of Ballroom Dancing. Ballroom dancing is as romantic as it is enjoyable. For years the world's ...

cinderelladanceschool.com/index.htm

[**Interactives . Elements of a Story . Cinderella**](#)

About this Interactive | Tips for Adults | Elements of a Story Site Map

www.learner.org/interactives/story/cinderella.html

[**Cinderella**](#)

I bought **Cinderella**, and it is running in German only. I have a Mac. **Cinderella** does not run on my Computer, although I know I have Java 2 installed on it [usually ...

cinderella.de/tiki-view_faq.php?faqId=1

Top 4 results

[bambi]

[Bambi - Wikipedia, the free encyclopedia](#)

[Plot](#) · [Cast](#) · [Production](#) · [Release](#) · [Reception](#) · [Legacy](#)

Bambi is a 1942 American animated film directed by David Hand (supervising a team of sequence directors), produced by Walt Disney and based on the book *Bambi, A ...*
en.wikipedia.org/wiki/Bambi

[Images of bambi](#)

See also: [Bambi 2 Vhs](#) · [Disney Images Of Bambi](#) · [Bambi And Feline](#)



[Bambi \(1942\) - IMDb](#)

Animation/Drama/Family · 70 min

Director: James Algar, Samuel Armstrong. . Actors: Hardie Albright: Adolescent Bambi · Stan Alexander: Young Flower · Bobette ...

www.imdb.com/title/tt0034492

IMDB

7.5/10

39,633 ratings

Top 3 results

The SETI-Capable BAMBI Radio Telescope

By Bob Lash and Mike Fremont

[ban]



Abstract

The design, construction, and initial observational results of a 4 GHz amateur radio telescope are described in this first report from Project BAMBI (Bob And Mike's Big Investment). The system is now operating continuously. The planned extension of the BAMBI project to amateur SETI is also discussed.

Introduction

A number of efforts are underway in the Search for Extraterrestrial Intelligence (SETI). We have been deeply interested in the search for some time, and have concluded that amateurs can in fact construct affordable systems with sensitivities comparable to professional all-sky search strategies even with antennas of limited aperture. We have also concluded that we can achieve a reasonably respectable frequency coverage of a search spectrum as well. We hope this project will encourage other amateurs to join in the search. Project BAMBI is divided into two phases:

Phase I: Standard Amateur Radio Astronomy

We have initially operated BAMBI as a total power receiver for several

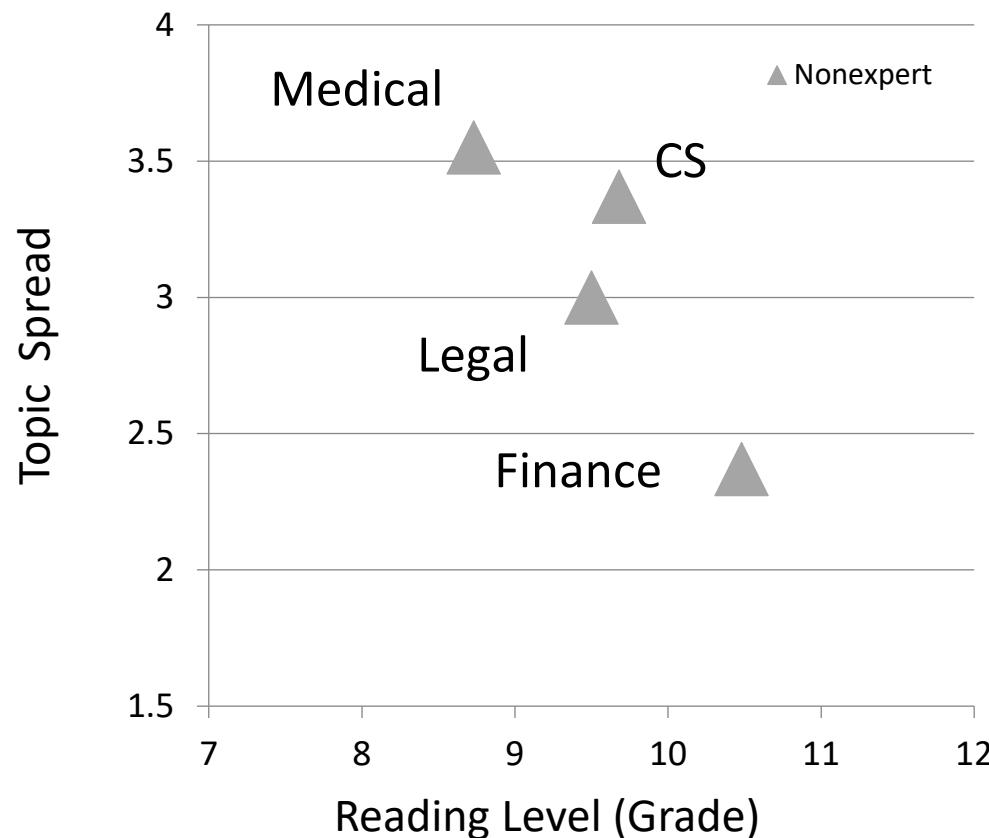
Adding reading level metadata to existing topic predictions helps model user and site expertise

[Kim, Collins-Thompson, Bennett, Dumais WSDM 2012]

- Four features:
 - Reading level
 1. Expected reading level $E(R)$
 2. Entropy $H(R)$ of reading level distribution
 - Topic
 3. Top- K ODP category predictions
 4. Entropy $H(T)$ of ODP category distribution
- Applications:
 - Classify expert vs non-expert sites and users
 - Better personalization features for search
 - Better click entropy prediction from topic/RL entropy
 - Spam/content farm detection
 - Detecting difficult tasks for users

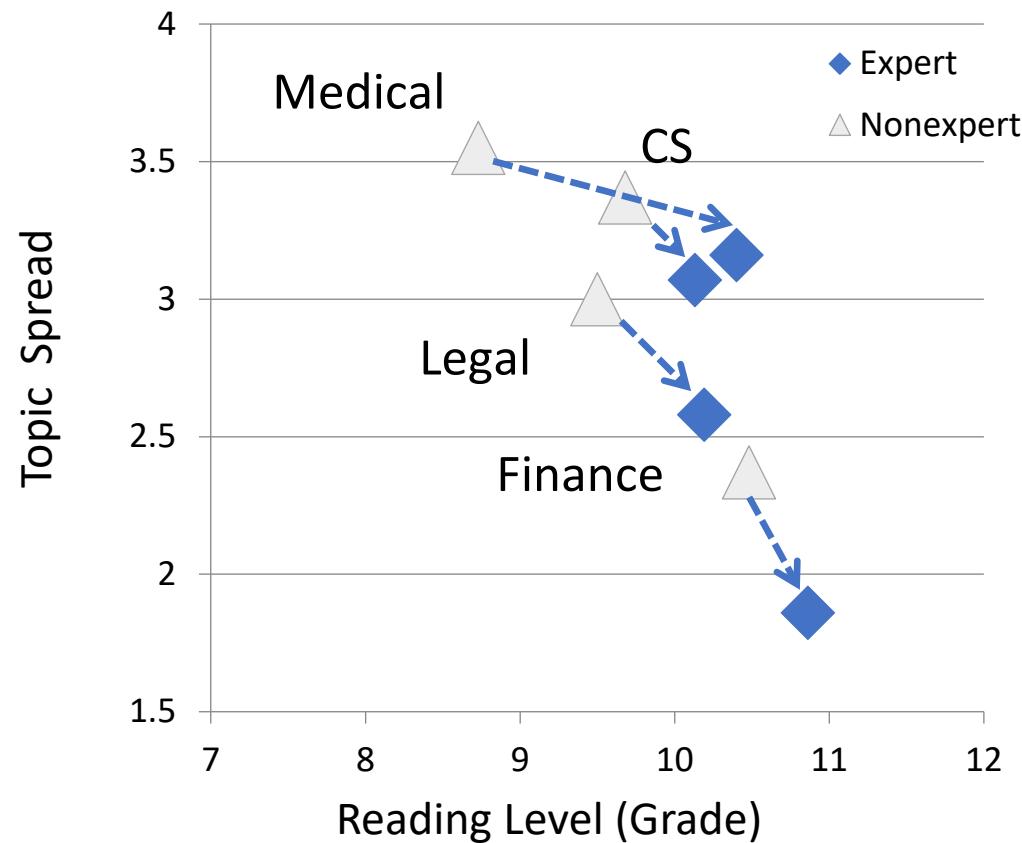
Reading level and topic entropy features can help separate expert from non-expert websites

[Kim, Collins-Thompson, Bennett, Dumais WSDM 2012]



Reading level and topic entropy features can help separate expert from non-expert websites

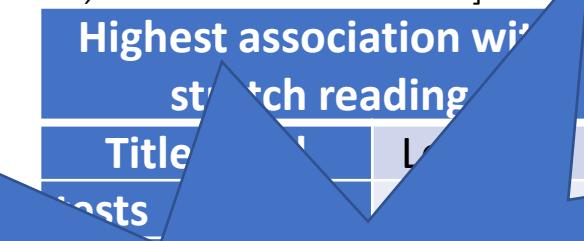
[Kim, Collins-Thompson, Bennett, Dumais WSDM 2012]



'Stretch' tasks: what are people searching for when they deviate from their typical reading level profile?

[Kim, Collins-Thompson, Bennett, Dumais: WSDM 2012]

From search sessions with unusually high deviation
> 4 levels deviation in user's average



Future work:

1. Identify & predict stretch tasks
2. How/when to provide support?
3. Find helpful alternatives



Medical tests
College entrance

Financial aid
Grievance forms
Job search

Based on two months of Bing traffic.

Adding reading level metadata improves click models for ranking.
Users can be misled by a mismatch between snippet difficulty & page difficulty
[Collins-Thompson, Bennett, White, de la Chica, Sontag, CIKM 2011]

Snippet Difficulty: Medium

ALL RESULTS 1-10 of 3,840,000 results - Advanced

Click!

Welcome to I REaR

INSECT REARING EDUCATION and RESEARCH PROGRAM
Department of Entomology, North Carolina State University
Allen Cohen, Program Coordinator

The I REaR program is dedicated to:

- The advancement of insect rearing as a science and technology**
- The education of students and rearing professionals in the most up-to-date rearing practices**
- The development of quality control and process control in rearing systems**
- Insectary problem solving and insectary techniques**
- The recognition that insect rearing must be done in a safe and green environment**

Retreat!!

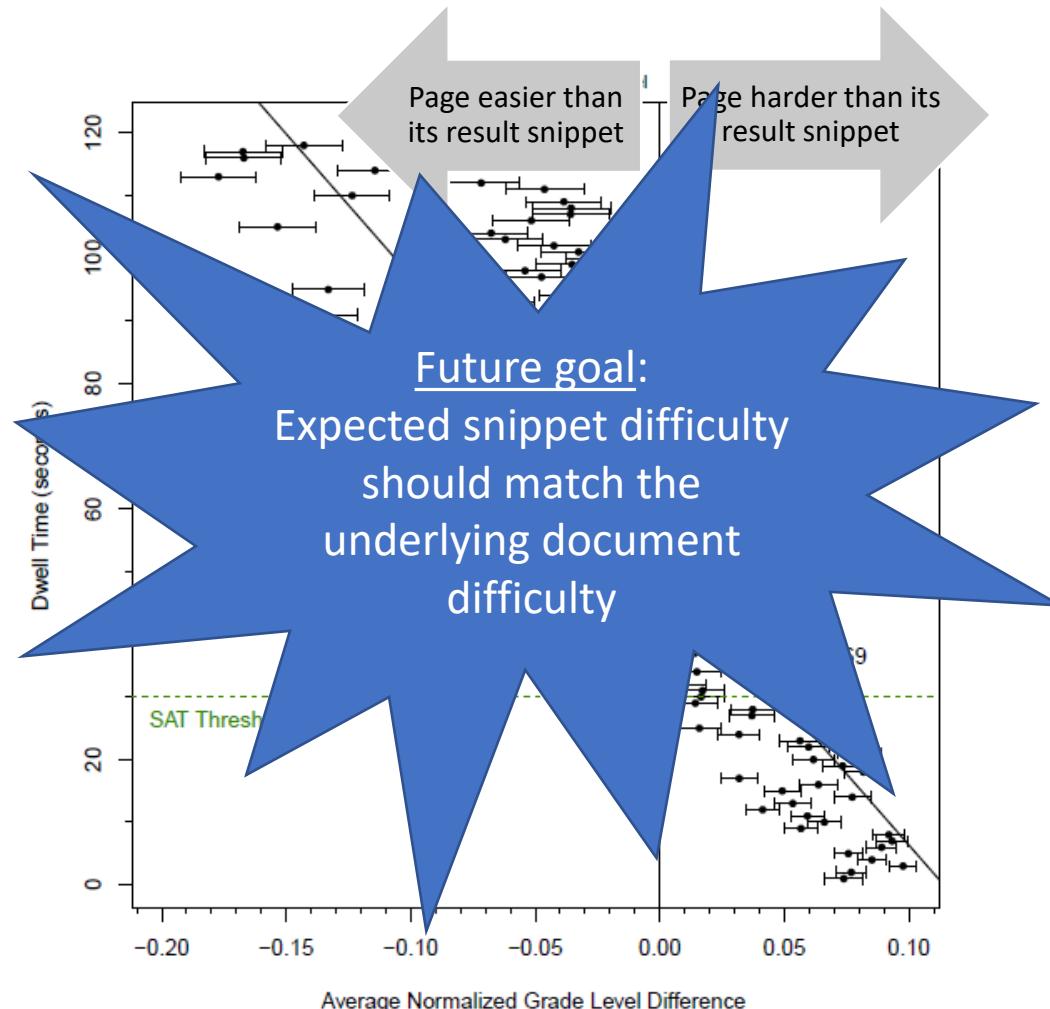
Insect Rearing Philosophy

- Rearing is treated as a science and technology**
- Knowing the insect thoroughly is essential to successful programs**
- Rearing is interdisciplinary**
- Good rearing systems must provide all the needs of insects (everything that insects need in nature must be provided in artificial rearing systems)**

Who Makes up the Insect Rearing Program at NCSU?

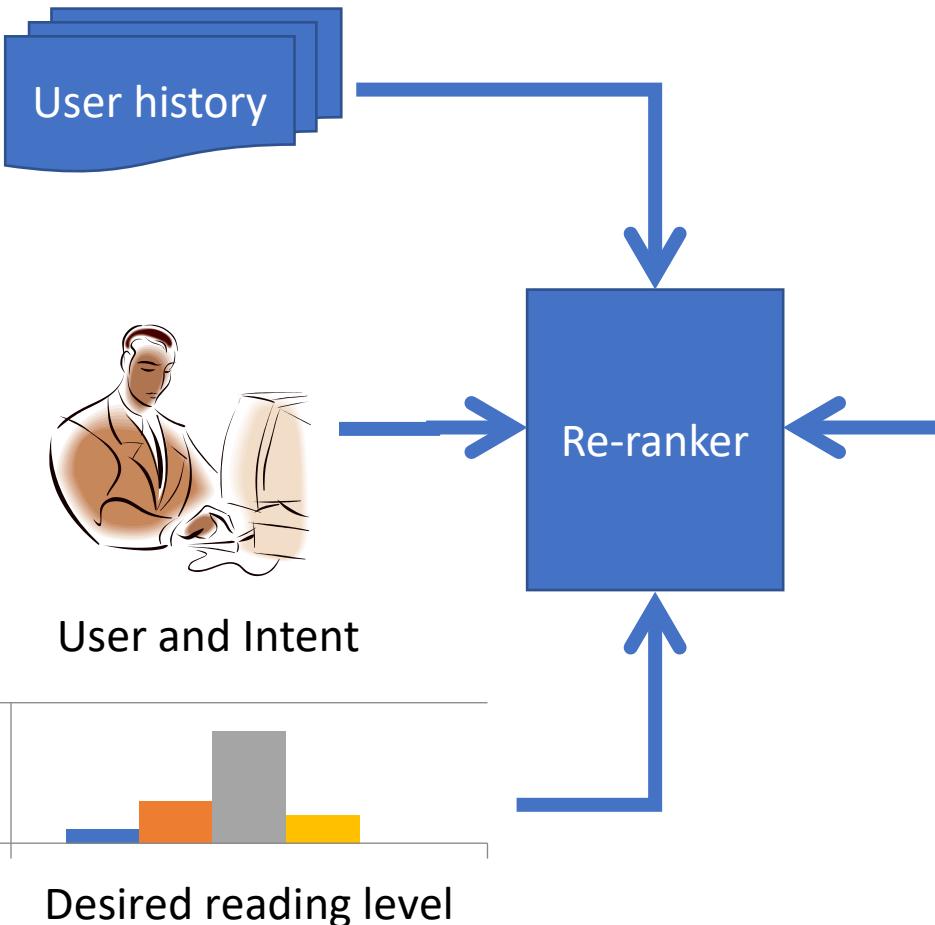
- The IREaR program is based in the NCSU Entomology Department, with extensive interactions and cooperation with other departments**

Users abandon pages faster when actual page is more difficult than the search result snippet suggested



Personalizing Web search by reading level

[Collins-Thompson, Bennett, White, de la Chica, Sontag, CIKM 2011]



Grasshopper Habitat and Grasshopper Diet

Grasshoppers live in fields, meadows and just about anywhere they can find generous amounts of food to eat. A grasshopper has a hard shell and a full grown grasshopper is about one and a half inches, being so small you would not think they would eat much - but you half think - they eat lots and lots - an average grasshopper can eat 16 time its own weight.

The grasshoppers favourite foods are grasses, leaves and cereal crops. One particular grasshopper - the Shorthorn grasshopper only eats plants, but it can go berserk and eat every plant in sight - makes you wonder where they put it all.



Grasshopper Behaviour

Insect Printouts

Firefly or Lightning Bug
Photinus pyralis

The Pyralis firefly (also known as the lightning bug) is a common firefly in North America. This partly nocturnal, luminescent beetle is the most common firefly in the USA.

The Firefly's Glow: At night, the very end (the last abdominal segment) of the firefly glows a bright yellow-green color. The firefly can control this glowing effect. The brightness of a single firefly is 1/40 of a candle. Fireflies use their glow to attract other fireflies. Males flash about every five seconds, females respond about every two seconds. This ability is harvested by the biochemical industry for the organic compound luciferin (which is the chemical the firefly uses for its bioluminescence).

Anatomy: This flying insect is about 0.75 inch (2 cm) long. It is mostly black, with two red spots on the head cover; the wing covers and head covers are lined in yellow. Like all insects, it has a hard exoskeleton, six jointed legs, two antennae, compound eyes, and a body divided into three parts (the head, thorax, and abdomen).

Diet: Both the adults and the larvae are **carnivores** (meat-eaters). They eat other insects (including other fireflies), insect larvae, and snails.

R&D

INSECT REARING
RESEARCH AND DEVELOPMENT
at NCSU

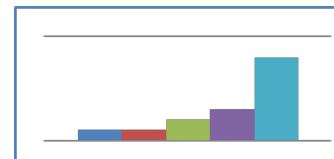
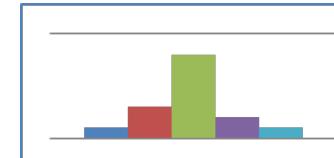
REARING RESEARCH ON DIET DEVELOPMENT AND ESTABLISHING NEW AND IMPROVED REARING SYSTEMS

- Development of artificial diets and rearing systems for many species of insects has been our specialty.
- We have a series of seminars that may be arranged upon request, data, including analysis of the natural foods, feeding behavior of the target insects, rearing, integrated-harvesting, and analysis of internal biology.

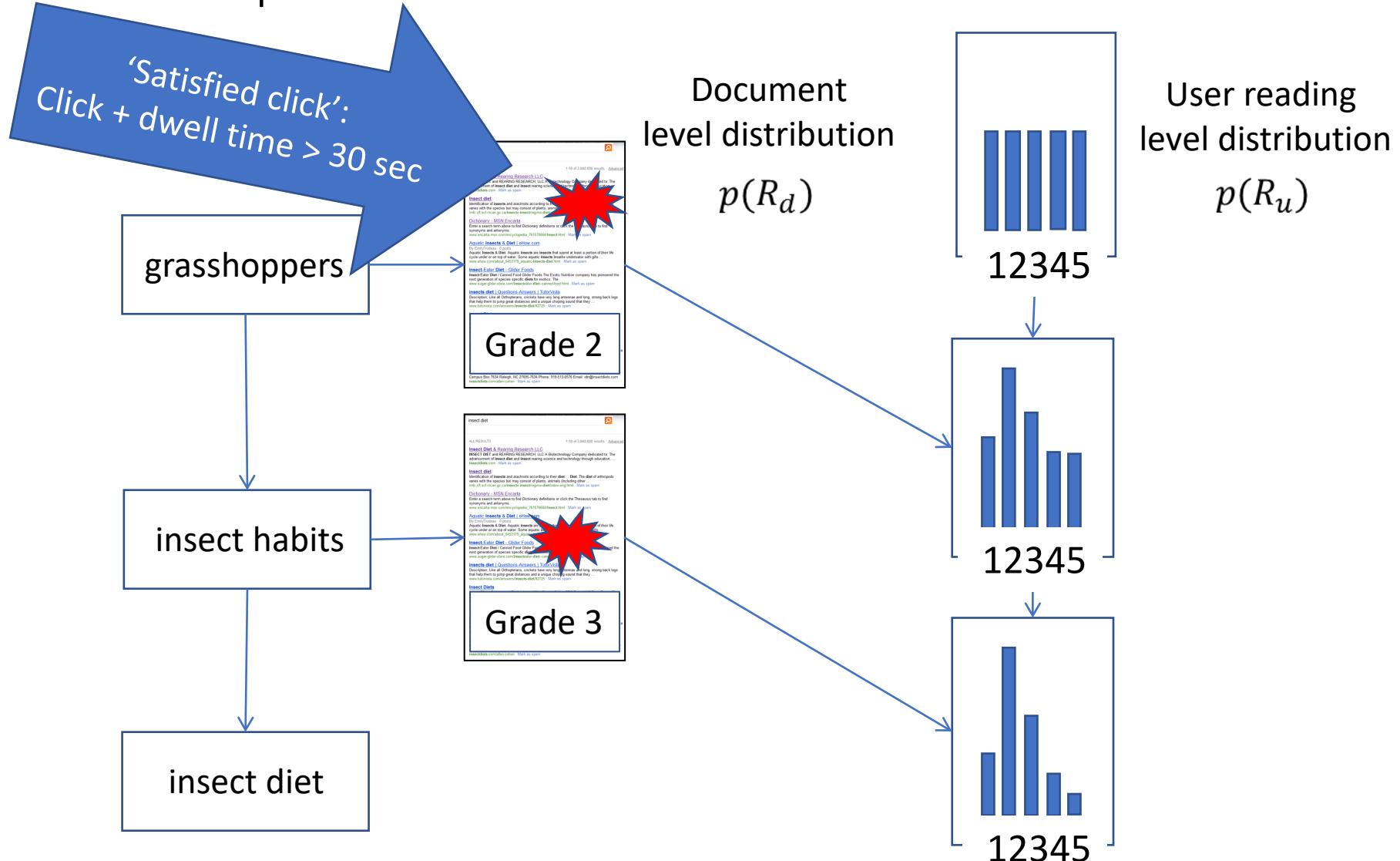
Some of our recent and current projects are below:



Content reading level



A simple generative user model combines reading levels of previous clicked documents in the session



Which features could help personalize?

- Content
 - Page reading level. (query-agnostic)
 - Result snippet reading level. (query-dependent)
- Query
 - Length in words, characters.
 - Query term reading level.
- User Session
 - All queries since last >30 min of inactivity.
 - Pages with satisfied clicks.
Assumption: user likes results \leq their desired level
- Interaction features
 - e.g. Snippet-Page difference in reading level.
- Variance of the above features.

What types of queries are helped most by reading level personalization?

Query subset	Num. queries	% Total	Gain
			Gain from general relevance features: + 0.5
			Gain from personalization features: + 0.7
			Point Gain in Mean Reciprocal Rank of last satisfied click
			relative to Bing production baseline.

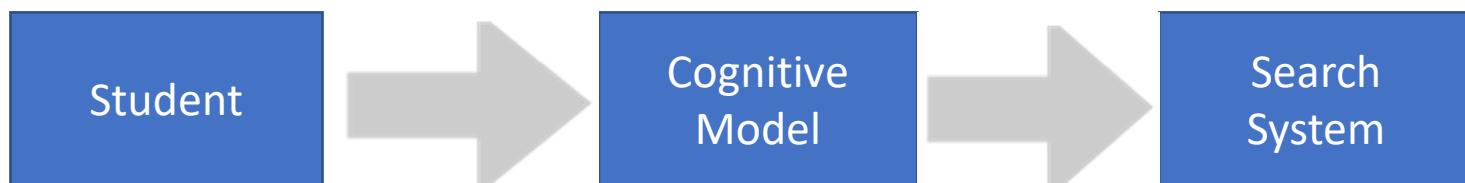
- Gain for all queries; varied with query subset
 - Any gain ≥ 1.0 over production is notable.
 - Science queries benefited most.
 - Net +1.6% of all queries improved at least one rank position in satisfied click
 - Large rank changes (> 5 positions) more than 70% likely to result in a win

Using reading level and word acquisition models to optimize search ranking for human learning
[Syed & Collins-Thompson, SIGIR 2017]

Key idea:

Optimize retrieval through the cognitive lens of the user

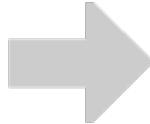
- Define representation of user's knowledge state
- Estimate prior and goal knowledge states
- Specify a cognitive model of how information affects the user's knowledge state.
 - i.e. how the user learns from the presented information



Vocabulary learning from context for the topic “igneous rock”

1 Determine Most Relevant Terms

Term	Weights
“rocks”	0.308
“igneous”	0.236
“magma”	0.139
“minerals”	0.081
“basalt”	0.056



https://en.wikipedia.org/wiki/Igneous_rock

Extrusive

Extrusive igneous rocks, also known as volcanic rocks, are formed at the crust's surface as a result of the partial melting of rocks within the mantle and crust. Extrusive igneous rocks cool and solidify quicker than intrusive igneous rocks. They are formed by the cooling of molten magma on the earth's surface. The magma, which is brought to the surface through fissures or volcanic eruptions, solidifies at a faster rate. Hence such rocks are smooth, crystalline and fine-grained. Basalt is a common extrusive igneous rock and forms lava flows, lava sheets and lava plateaus. Some kinds of basalt solidify to form long columnar columns.



Extrusive igneous rock is made from lava released by volcanoes



3 Student takes test after reading docs

Exam

1	_____	_____	_____	✓
2	_____	_____	_____	✗
3	_____	_____	_____	✓



2 Provide documents covering terms

⋮

<https://uwaterloo.ca/wat-on-earth/news/basic-introduction-rocks-part-i-igneous-rocks>

as plates, cups and saucers (pottery, ceramic and china), counter tops and of various origins. Rocks are aggregates of different mineral grains and can be major families or rock groupings.

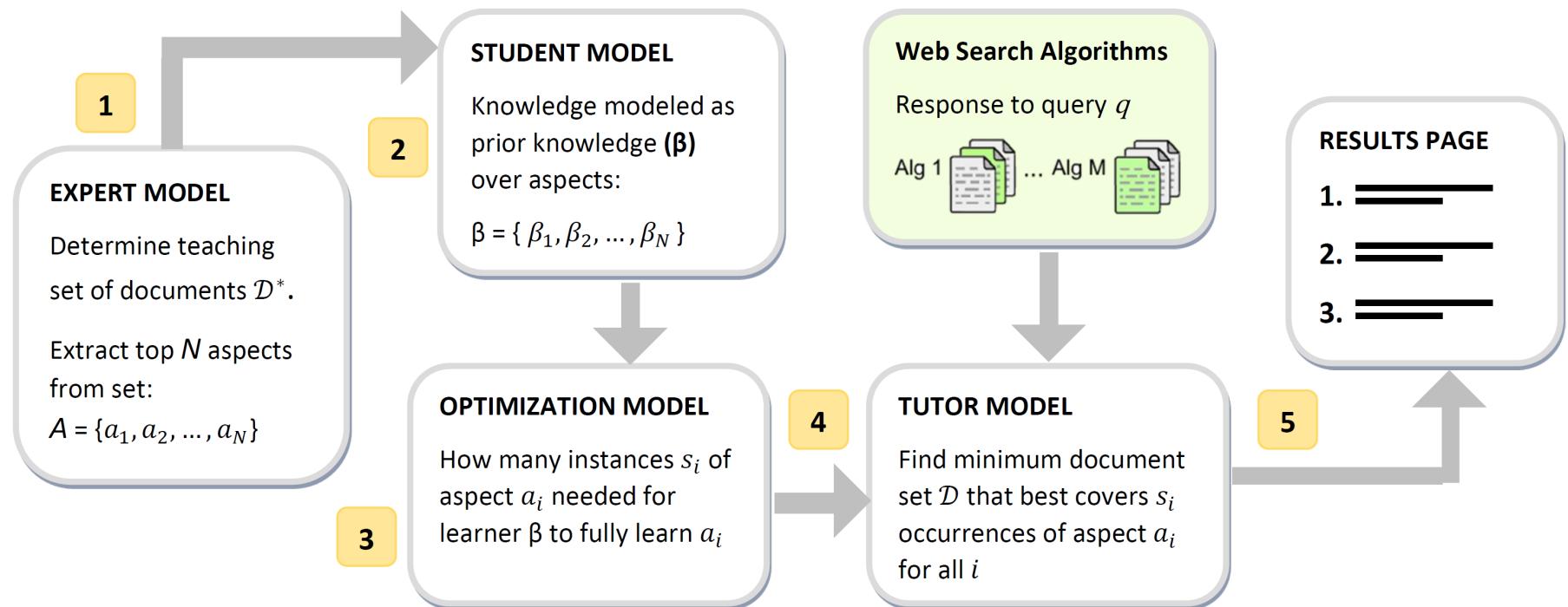
First are the Igneous (or "fire-formed") Rocks, usually created by outpouring from volcanoes or by cooling deep under the crust. Ultimately, even deeply buried rocks can be exposed to surface weathering and break down into their constituent minerals. These minerals can be removed as sediment and are transported by gravity, wind, ice and water, where they accumulate normally as marine sediments. The sediments can then be

[Syed and Collins-Thompson, 2017]

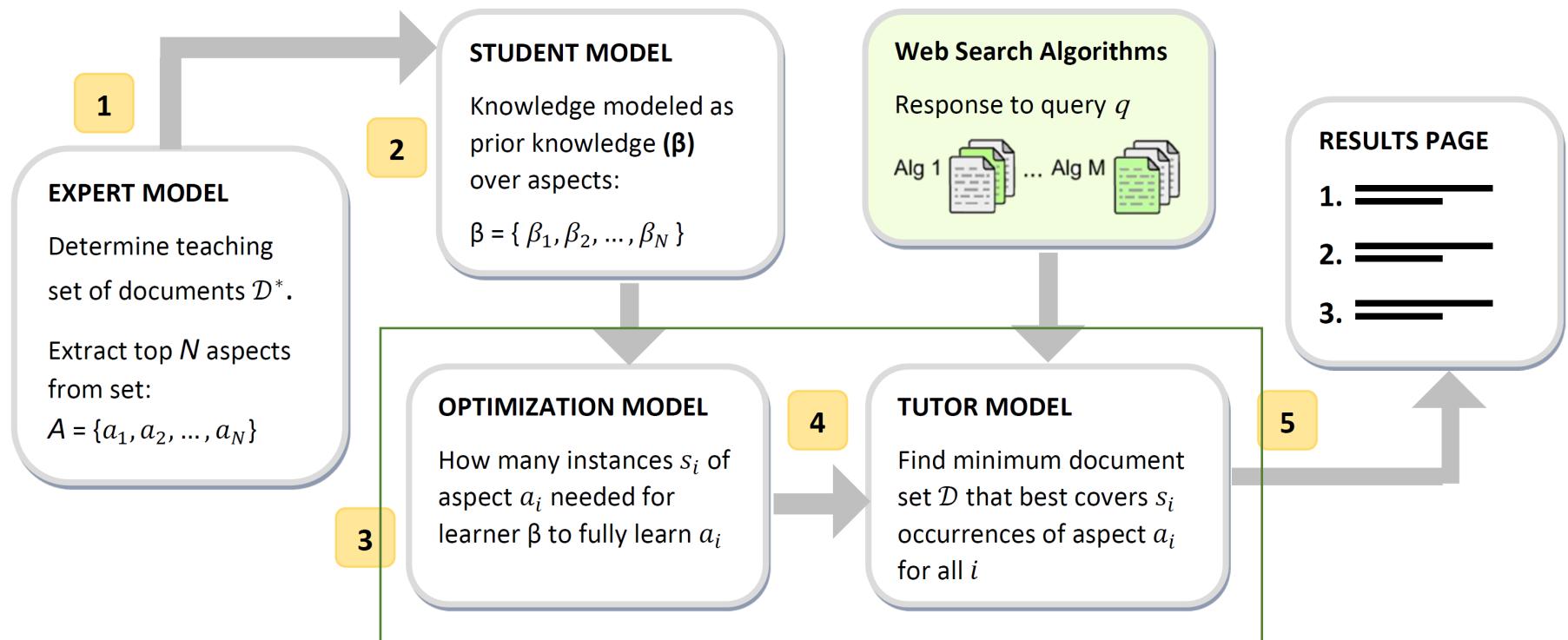
KDD Doc Intelligence Aug 15, 2021

© 2001-2021 K. Collins-Thompson

Goal: Search algorithms that support effective vocabulary learning from context



Our SIGIR 2017 paper addressed the personalized retrieval problem in steps 3 and 4.



Optimally connecting users with information for personalized learning

What should a retrieval objective for learning look like? Find information that...

1. *Advances an individual user's learning progress toward a specific goal.*

2. *Minimizes or reduces effort toward that goal.*

Overall retrieval optimization problem:
find an optimal set of documents D

$$\operatorname{argmax}_D U(D) = H(D) - \lambda \cdot E(D)$$



Searching efficiently
for optimal set D can be hard...

Search for an approximately optimal document set D into two steps

- (a) How many total exposures S_k for each word are needed for each item k to efficiently maximize the learning outcome?
- (b) Find a 'good' set of documents that exposes the user to these optimal S_k per-word exposures in context.
- (c) User effort modeled by length and reading difficulty of the material.

Personalized ranking gave significantly better learning outcomes than generic search (and best overall)

Measure	Web (n=290)	Non-Personalized (n=290)	Personalized (n=283)	p-value
Absolute Learning Gains	1.72	1.83	1.98	p=.046
Realized Potential Learning	0.38	0.43	0.47	p=.008
Learning Gains/ 1000 Words	0.11	0.25	0.35	p<.001

Type of Test: Kruskal Wallis H Test - Omnibus

Personalization helps long-term learning, confirmed also in:
CHIIR 2018: Exploring Document Retrieval Features Associated with Improved Short- and Long-term Vocabulary Learning Outcomes
[Syed and Collins-Thompson, 2018]

Adaptive Learning for Reading Comprehension: Gaze classifier tracks reading fixations, generates adjunct questions for the reader

[Syed, Collins-Thompson, Bennett, Teng, Williams, Tay, Iqbal. WWW 2020]

Adaptive Learning

Article Apollo 11

WIKIPEDIA The Free Encyclopedia

Article Talk

From Wikipedia, the free encyclopedia

This article is about the 1969 manned lunar mission. For other uses, see [Apollo 11 \(disambiguation\)](#).

Apollo 11 was the spaceflight that landed the first two people on the Moon. Mission commander Neil Armstrong and pilot Buzz Aldrin, both American, landed the lunar module *Eagle* on July 20, 1969, at 20:17 UTC. Armstrong became the first person to step onto the lunar surface six hours after landing on July 21 at 02:56:15 UTC; Aldrin joined him about 20 minutes later. They spent about two and a quarter hours together outside the spacecraft, and collected 47.5 pounds (21.5 kg) of lunar material to bring back to Earth. Michael Collins piloted the command module *Columbia* alone in lunar orbit while they were on the Moon's surface. Armstrong and Aldrin spent 21.5 hours on the lunar surface before rejoining Columbia in lunar orbit.

Apollo 11 was launched by a Saturn V rocket from Kennedy Space Center on Merritt Island, Florida, on July 16 at 9:32 am EDT (13:32 UTC) and was the first manned mission of NASA's Apollo program. The Apollo spacecraft had three parts: a command module (CM) with a cabin for the three astronauts, and the only part that returned to Earth; a service module (SM), which supported the command module with propulsion, electrical power, oxygen, and water; and a lunar module (LM) that had two stages – a descent stage for landing on the Moon and an ascent stage to place the astronauts back into lunar orbit.

After being sent to the Moon by the Saturn V's third stage, the astronauts separated the spacecraft from it and traveled for three days until they entered into lunar orbit. Armstrong and Aldrin then moved into the lunar module *Eagle* and landed in the Sea of Tranquility. The astronauts used *Eagle*'s upper stage to lift off from the lunar surface and return Collins in the command module. They jettisoned *Eagle* before the descent stage's engines blasted them out of lunar orbit on a trajectory back to Earth. They returned to Earth and splashed down in the Pacific Ocean on July 24 after more than eight days in space.

The landing was broadcast on live TV to a worldwide audience. Armstrong stepped onto the lunar surface and described the event as "one small step for [a] man, one giant leap for mankind." Apollo 11 effectively ended the Space Race and fulfilled a national goal proposed in 1961 by U.S. President John F. Kennedy: "before this decade is out, of landing a man on the Moon and returning him safely to the Earth."^[8]

Apollo 11

Neil Armstrong during the deployment of the United States flag during the EVA on the lunar surface.

Mission type Manned lunar landing

Launch date 1969-07-16 00:00:00 UTC

Orbital period 2 hours, 3 minutes, 56 seconds

COSPAR ID CSM: 1969-059A

LM: 1969-059C

SATCAT no. CSM: 4039

LM: 4041

Mission duration 8 days, 3 hours, 18 minutes, 35 seconds

Spacecraft properties

Spacecraft Apollo CSM-107

Apollo LM-5

Manufacturer CSM: North American

Get Question

Check Answer

Contents [hide]

1 Framework

1.1 Crew

1.2 Backup crew

1.3 Support crew

1.4 Flight directors

1.5 Call signs

1.6 Insignia

1.7 Mementos

2 Mission highlights

Attention is approximately quantified by Normalized Number of Fixations (per word) in content window: Fixation types: **skimming**, **reading**, **regression**

Adaptive Learning

Article Apollo 11

Show Reading Overlay Clear Configuration

Fixation Type	Text Segment	Normalized Number of Fixations (NNF)
SKIM	In addition to millions of people crowding highways and beaches near the launch site, millions watched the event on television, with NASA Chief of Public Information Jack King providing commentary. President Richard M. Nixon viewed the proceedings from the Oval Office of the White House. ^[1]	0.22
SKIM	A Saturn V launched Apollo 11 from Launch Pad 39A, part of the Launch Complex 39 site at the Kennedy Space Center on July 16, 1969, at 13:32:00 UTC (9:32:00 a.m. EDT local time). It entered orbit, at an altitude of 100.4 nautical miles (185.9 km) by 98.9 nautical miles (183.2 km), twelve minutes later. After one and a half orbits, the S-IVB third-stage engine pushed the spacecraft onto its trajectory toward the Moon with the trans-lunar injection (TLI) burn at 16:22:13 UTC. About 30 minutes later, the transposition, docking, and extraction maneuver was performed: this involved separating the Apollo Command/Service Module (CSM) from the spent rocket stage, turning around, and docking with the Lunar Module still attached to the stage. After the Lunar Module was extracted, the combined spacecraft headed for the Moon, while the rocket stage flew on a trajectory past the Moon and into orbit around the Sun. ^[2]	0.05
SKIM	On July 19 at 17:21:50 UTC, Apollo 11 passed behind the Moon and fired its service propulsion engine to enter lunar orbit. In the thirty orbits ^[2] that followed, the crew saw passing views of their landing site in the southern Sea of Tranquility (Mare Tranquillitatis) about 12 miles (19 km) southwest of the crater Sabine D (0.67406N, 23.47297E). The landing site was selected in part because it had been characterized as relatively flat and smooth by the automated Ranger 8 and Surveyor 5 landers along with the Lunar Orbiter mapping spacecraft and unlikely to present major landing or extravehicular activity (EVA) challenges. ^[2]	0.08
SKIM	Lunar descent	0.13
SKIM	On July 20, 1969, the Lunar Module <i>Eagle</i> separated from the Command Module <i>Columbia</i> . Collins, alone aboard <i>Columbia</i> , inspected <i>Eagle</i> as it pirouetted before him to ensure the craft was not damaged.	0.06
READ	As the descent began, Armstrong and Aldrin found that they were passing landmarks on the surface four seconds early and reported that they were "long": they would land miles west of their target point.	1.31
READ	Five minutes into the descent, at 6,000 feet (1,800 m) above the surface of the Moon, the LM navigation and guidance computer distracted the crew with the firing of several unexpected "1202" and "1201" program alarms. Inside Mission Control Center in Houston, Texas, computer engineer Jack Garman told guidance officer Steve Bales it was safe to continue the descent, and this was relayed to the crew. The program alarms indicated "executive overflows", meaning the guidance computer could not complete all of its tasks in real time and had to postpone some of them. ^[24]	0.06
READ	Due to an error in the checklist manual, the rendezvous radar switch was placed in the wrong position. This caused it to send erroneous signals to the computer. The result was that the computer was being asked to perform all of its normal functions for landing while receiving an extra load of spurious data which used up 15% of its time. The computer (or rather the software in it) was smart enough to recognize that it was being asked to perform more tasks than it should be performing. It then sent out an alarm, which meant to the astronaut, I'm overloaded with more tasks than I should be doing at this time and I'm going to keep only the more important tasks, i.e., the ones needed for landing ... Actually, the computer was programmed to do more than recognize error conditions. A complete set of recovery programs was incorporated into the software. The software's action, in this case, was to eliminate lower priority tasks and re-establish the more important ones ... If the computer hadn't recognized this problem and taken recovery action, I doubt if Apollo 11 would have been the successful moon landing it was. ^{[25][26]}	0.06
—Letter from Margaret H. Hamilton, Director of Apollo Flight Computer Programming MIT Draper Laboratory, Cambridge, Massachusetts ^[27] , titled "Computer Got Loaded", published in <i>Datamation</i> , March 1, 1971		0.06
Where was the landing broadcast?		Get Question Difficult Check Answer

Regression fixations indicate material that a user had to go back and re-read

Easy questions are generated from content read more closely

Adaptive Learning

Article Apollo 11

WIKIPEDIA The Free Encyclopedia

Article Talk

Read View source View history Search Wikipedia

Apollo 11

From Wikipedia, the free encyclopedia

This article is about the 1969 manned lunar mission. For other uses, see [Apollo 11 \(disambiguation\)](#).

Apollo 11 was the spaceflight that landed the first two people on the Moon. Mission commander Neil Armstrong and pilot Buzz Aldrin, both American, landed the lunar module *Eagle* on July 20, 1969, at 20:17 UTC. Armstrong became the first person to step onto the lunar surface six hours after landing on July 21 at 02:56:15 UTC; Aldrin joined him about 20 minutes later. They spent about two and a quarter hours together outside the spacecraft, and collected 47.5 pounds (21.5 kg) of lunar material to bring back to Earth. Michael Collins piloted the command module *Columbia* alone in lunar orbit while they were on the Moon's surface. Armstrong and Aldrin spent 21.5 hours on the lunar surface before rejoining Columbia in lunar orbit.

Apollo 11 was launched by a Saturn V rocket from Kennedy Space Center on Merritt Island, Florida, on July 16 at 9:32 am EDT (13:32 UTC) and was the fifth manned mission of NASA's Apollo program. The Apollo spacecraft had three parts: a command module (CM) with a cabin for the three astronauts, and the only part that returned to Earth; a service module (SM), which supported the command module with propulsion, electrical power, oxygen, and water; and a lunar module (LM) that had two stages – a descent stage for landing on the Moon and an ascent stage to place the astronauts back into lunar orbit.

After being sent to the Moon by the Saturn V's third stage, the astronauts separated the spacecraft from it and traveled for three days until they entered into lunar orbit. Armstrong and Aldrin then moved into the lunar module *Eagle* and landed in the Sea of Tranquility. The astronauts used *Eagle*'s upper stage to lift off from the lunar surface and return Collins in the command module. They jettisoned *Eagle* before the descent stage's engines blasted them out of lunar orbit on a trajectory back to Earth. They returned to Earth and splashed down in the Pacific Ocean on July 24 after more than eight days in space.

The landing was broadcast on live TV to a worldwide audience. Armstrong stepped onto the lunar surface and described the event as "one small step for [a] man, one giant leap for mankind." Apollo 11 effectively ended the Space Race and fulfilled a national goal proposed in 1961 by U.S. President John F. Kennedy: "before this decade is out, of landing a man on the Moon and returning him safely to the Earth."^[8]

Contents [hide]

- 1 Framework
 - 1.1 Crew
 - 1.2 Backup crew
 - 1.3 Support crew
 - 1.4 Flight directors
 - 1.5 Call signs
 - 1.6 Insignia
 - 1.7 Mementos
- 2 Mission highlights

What was the name of the mission that landed the first two people on the moon?

Get Question Easy Check Answer

Apollo 11

Neil Armstrong during the deployment of the United States flag during the EVA on the lunar surface.

Mission type Manned lunar landing

Launch date 1969-07-16 00:00:00 UTC

Orbiter CSM-107, LM-04

COSPAR ID 20350

LM: 1969-059C

SATCAT no. 4039

LM: 4041

Mission duration 8 days, 3 hours, 18 minutes, 35 seconds

Spacecraft properties

Spacecraft Apollo CSM-107

Apollo LM-5

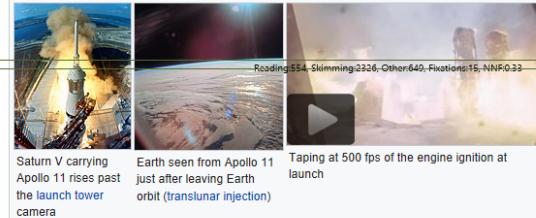
Manufacturer CSM: North American

Automatic question: What was the name of the mission...

Launch and flight to lunar orbit

In addition to millions of people crowding highways and beaches near the launch site, millions watched the event on television, with NASA Chief of Public Information Jack King providing commentary. President Richard M. Nixon viewed the proceedings from the Oval Office of the White House. [citation needed]

A Saturn V launched Apollo 11 from Launch Pad 39A, part of the Launch Complex 39 site at the Kennedy Space Center on July 16, 1969, at 13:32:00 UTC (9:32:00 a.m. EDT local time). It entered Earth orbit, at an altitude of 100.4 nautical miles (185.9 km) by 98.9 nautical miles (183.2 km), twelve minutes later. After one and a half orbits, the S-IVB third-stage engine pushed the spacecraft onto its trajectory toward the Moon with the trans-lunar injection (TLI) burn at 16:22:13 UTC. About 30 minutes later, the transposition, docking, and extraction maneuver was performed; this involved separating the Apollo Command/Service Module (CSM) from the spent rocket stage, turning around, and docking with the Lunar Module still attached to the stage. After the Lunar Module was extracted, the combined spacecraft headed for the Moon, while the rocket stage flew on a trajectory past the Moon and into orbit around the Sun.^[3]



On July 19 at 17:21:50 UTC, Apollo 11 passed behind the Moon and fired its service propulsion engine to enter lunar orbit. In the thirty orbits^[22] that followed, the crew saw passing views of their landing site in the southern Sea of Tranquility (Mare Tranquillitatis) about 12 miles (19 km) southwest of the crater Sabine D (0.67408N, 23.47297E).^[45] The landing site was selected in part because it had been characterized as relatively flat and smooth by the automated *Ranger 8* and *Surveyor 5* landers along with the *Lunar Orbiter* mapping spacecraft and unlikely to present major landing or extravehicular activity (EVA) challenges.^[23]

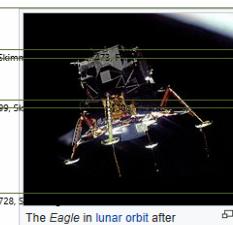
Reading 0, Skimming 1885, Other 1551, Fixations 8, NNF 0.08

Lunar descent

On July 20, 1969, the Lunar Module *Eagle* separated from the Command Module *Columbia*. Collins, alone aboard *Columbia*, inspected *Eagle* as it pirouetted before him to ensure the craft was not damaged.

As the descent began, Armstrong and Aldrin found that they were passing landmarks on the surface four seconds early and reported that they were "long"; they would land miles west of their target point.

Five minutes into the descent burn, and 6,000 feet (1,800 m) above the surface of the Moon, the LM navigation and guidance computer displayed several unexpected "1202" and "1201" program alarms. Inside Mission Control Center in Houston, Texas, computer engineer Jack Garman told guidance officer Steve Bales it was safe to continue the descent and this was relayed to the crew. The program alarms indicated "executive overflows", meaning the guidance computer could not complete all of its tasks in real time and had to postpone some of them.^[24]



The *Eagle* in lunar orbit after separating from *Columbia*

Due to an error in the checklist manual, the rendezvous radar switch was placed in the wrong position. This caused it to send erroneous signals to the computer. The result was that the computer was being asked to perform all of its normal functions for landing while receiving an extra load of spurious data which used up 15% of its time. The computer (or rather the software in it) was smart enough to recognize that it was being asked to perform more tasks than it should be performing. It then sent out an alarm, which meant to the astronaut, *I'm overloaded with more tasks than I should be doing at this time and I'm going to keep only the more important tasks; i.e., the ones needed for landing ...* Actually, the computer was programmed to do more than recognize error conditions. A complete set of recovery programs was incorporated into the software. The software's action, in this case, was to eliminate lower priority tasks and re-establish the more important ones ... If the computer hadn't recognized this problem and taken recovery action, I doubt if Apollo 11 would have been the successful moon landing it was.^{[25][6]}

— Letter from Margaret H. Hamilton, Director of Apollo Flight Computer Programming MIT Draper Laboratory, Cambridge, Massachusetts^[25], titled "Computer Got I Loaded", published in *Datamation*, March 1, 1971



What indicated the guidance computer could not complete all of its tasks in real time?

[Get Question](#)

Easy

[Check Answer](#)

Automatic question: What indicated the guidance computer could not complete all of its tasks in real time?

Difficult questions are generated from content that was skimmed

Adaptive Learning

Article Apollo 11

WIKIPEDIA The Free Encyclopedia

Main page Contents Featured content Current events Random article Donate to Wikipedia Wikipedia store

Interaction Help About Wikipedia Community portal Recent changes Contact page

Tools What links here Related changes Upload file Special pages Permanent link Page information Wikimedia item Cite this page

Print/export Create a book Download as PDF Printable version

In other projects Wikimedia Commons

Where was the landing broadcast?

Get Question Difficult Check Answer

Apollo 11

From Wikipedia, the free encyclopedia

This article is about the 1969 manned lunar mission. For other uses, see [Apollo 11 \(disambiguation\)](#).

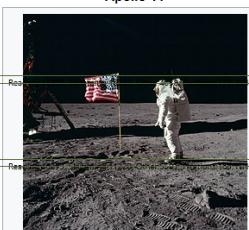
Apollo 11 was the spaceflight that landed the first two people on the Moon. Mission commander Neil Armstrong and pilot Buzz Aldrin, both American, landed the lunar module *Eagle* on July 20, 1969, at 20:17 UTC. Armstrong became the first person to step onto the lunar surface six hours after landing on July 21 at 02:56:15 UTC; Aldrin joined him about 20 minutes later. They spent about two and a quarter hours together outside the spacecraft, and collected 47.5 pounds (21.5 kg) of lunar material to bring back to Earth. Michael Collins piloted the command module *Columbia* alone in lunar orbit while they were on the Moon's surface. Armstrong and Aldrin spent 21.5 hours on the lunar surface before rejoining Columbia in lunar orbit.

Apollo 11 was launched by a Saturn V rocket from Kennedy Space Center on Merritt Island, Florida, on July 16 at 9:32 am EDT (13:32 UTC) and was the first manned mission of NASA's Apollo program. The Apollo spacecraft had three parts: a command module (CM) with a cabin for the three astronauts, and the only part that returned to Earth; a service module (SM), which supported the command module with propulsion, electrical power, oxygen, and water; and a lunar module (LM) that had two stages – a descent stage for landing on the Moon and an ascent stage to place the astronauts back into lunar orbit.

After being sent to the Moon by the Saturn V's third stage, the astronauts separated the spacecraft from it and traveled for three days until they entered into lunar orbit. Armstrong and Aldrin then moved into the lunar module *Eagle* and landed in the Sea of Tranquility. The astronauts used *Eagle*'s upper stage to lift off from the lunar surface and return Collins in the command module. They jettisoned *Eagle* before the descent stage's engines blasted them out of lunar orbit on a trajectory back to Earth. They returned to Earth and splashed down in the Pacific Ocean on July 24 after more than eight days in space.

The landing was broadcast on live TV to a worldwide audience. Armstrong stepped onto the lunar surface and described the event as "one small step for [a] man, one giant leap for mankind." Apollo 11 effectively ended the Space Race and fulfilled a national goal proposed in 1961 by U.S. President John F. Kennedy: "before this decade is out, of landing a man on the Moon and returning him safely to the Earth."^[8]

Apollo 11



During the EVA on the lunar surface, the television camera was deployed and transmitted live back to Earth.

Mission type	Manned lunar landing
Operator	NASA, Skimming 2118, CSM-21945, Fixations:26, NNF:0.34
COSPAR ID	CSM: 1969-059A LM: 1969-059C
SATCAT no.	CSM: 4039 LM: 4041
Mission duration	8 days, 3 hours, 18 minutes, 35 seconds
Spacecraft properties	
Spacecraft	Apollo CSM-107 Apollo LM-5
Manufacturer	CSM: North American

Automatic question: Where was the landing broadcast?

Most significant learning effects were detected in the long-term condition, especially for low-knowledge learners

[Syed, Collins-Thompson, Bennett, Teng, Williams, Tay, Iqbal. WWW 2020]

Result	Short-term learning	Long-term retention
Adjunct Questions improved grades better than QNone (Section 5.2)	No	Yes (for low-knowledge learners)
QAuto performed comparable to QHuman (Section 5.3)	Yes	Yes
Synthesis question affected grades (Section 5.4)	No	No
Focus-based question selection improved grades (Section 5.5)	No	Yes
Gaze behavior was different for those who would answer questions correctly (Section 6.4)	Yes (for low-knowledge learners)	Yes (for low-knowledge learners)

- Key takeaway: always have a delayed post-test!

Selected next steps

1. Domain-specific, personalized measures of content difficulty.
2. Modeling desirable difficulty in content.
3. Search and recommendation for educational videos.
 - a. Multi-modal documents: transcript, slides, lecture images, audio, supporting materials...
 - b. Local difficulty and complexity in lecture segments
 - c. Lots of business applications also: connect with OCR
4. Deep learning models of contextual informativeness.
 - a. Finding supportive text for learning new concepts.
 - b. Curriculum learning for *machine* readers.
 - c. Explainable models.

Our basic statistical readability models are available via REST API for non-commercial research use:

api.dscovar.org

Long-term: Modeling content difficulty for better support of human learning

1. Develop richer representations of users and content to support learning.
2. Integrate rich models of human learning into search and recommender algorithm objectives, features, and evaluation.
3. Continue developing reliable automated methods for explicit and implicit assessment of difficulty and learning during interaction.
4. Aim for robust long-term retention, not just short-term learning.

Contact: Kevyn Collins-Thompson

kevynct@umich.edu

Readability API: api.dscovar.org

Homepage: www.umich.edu/~kevynct

This research supported in part by:



This research supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305A140647 to the University of Michigan. The opinions expressed are those of the author and do not represent views of the Institute or the U.S. Department of Education.