Research Fast and Slow

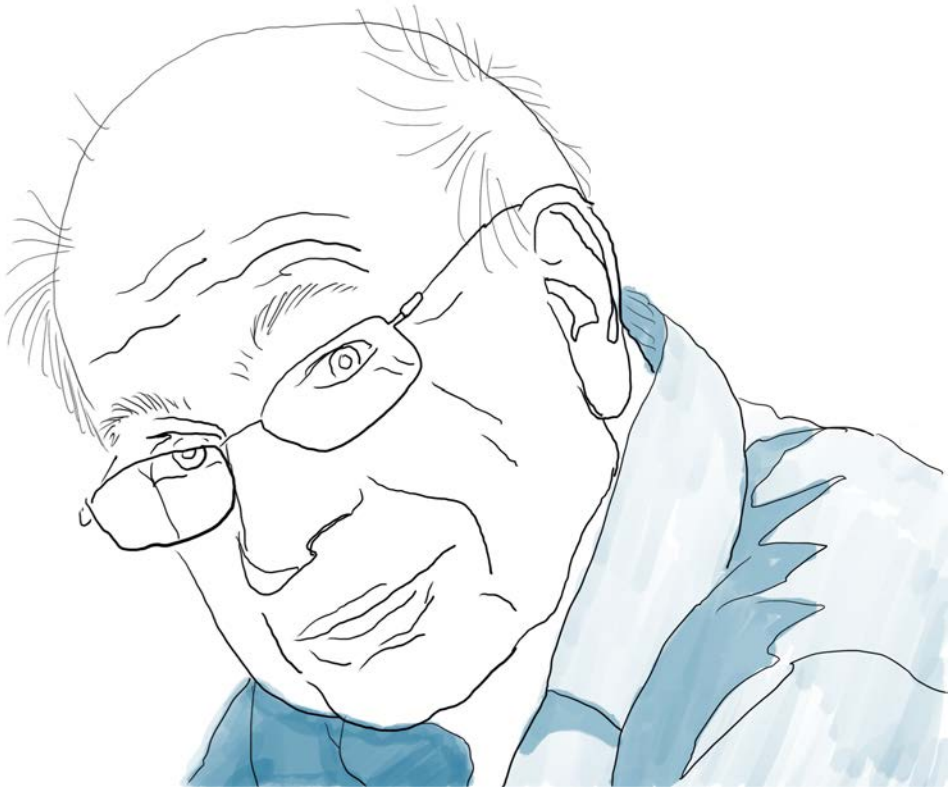
Min-Yen Kan

National University of Singapore

Slides available at http://bit.ly/kan-coling18

Slides: http://bit.ly/kan-coling18

Fast and Slow



Daniel Kahneman

System 1	System 2
Fast	Slow
Automatic	Controlled
Intuitive	Analytical
Parallel	Serial
Associative	Logical

The Neural Net – judgements in 1 second

Andrew Ng

System 1	System 2
Fast	Slow
Automatic	Controlled
Intuitive	Analytical
Parallel	Serial
Associative	Logical

To think about: what's the loss function of research?

The Age of Accelerations

Thomas Friedman

Thank You for Being Late

Kurzweil's "Second half of the chessboard"

His three accelerations

- Moore's law
- Globalization
- Mother Nature

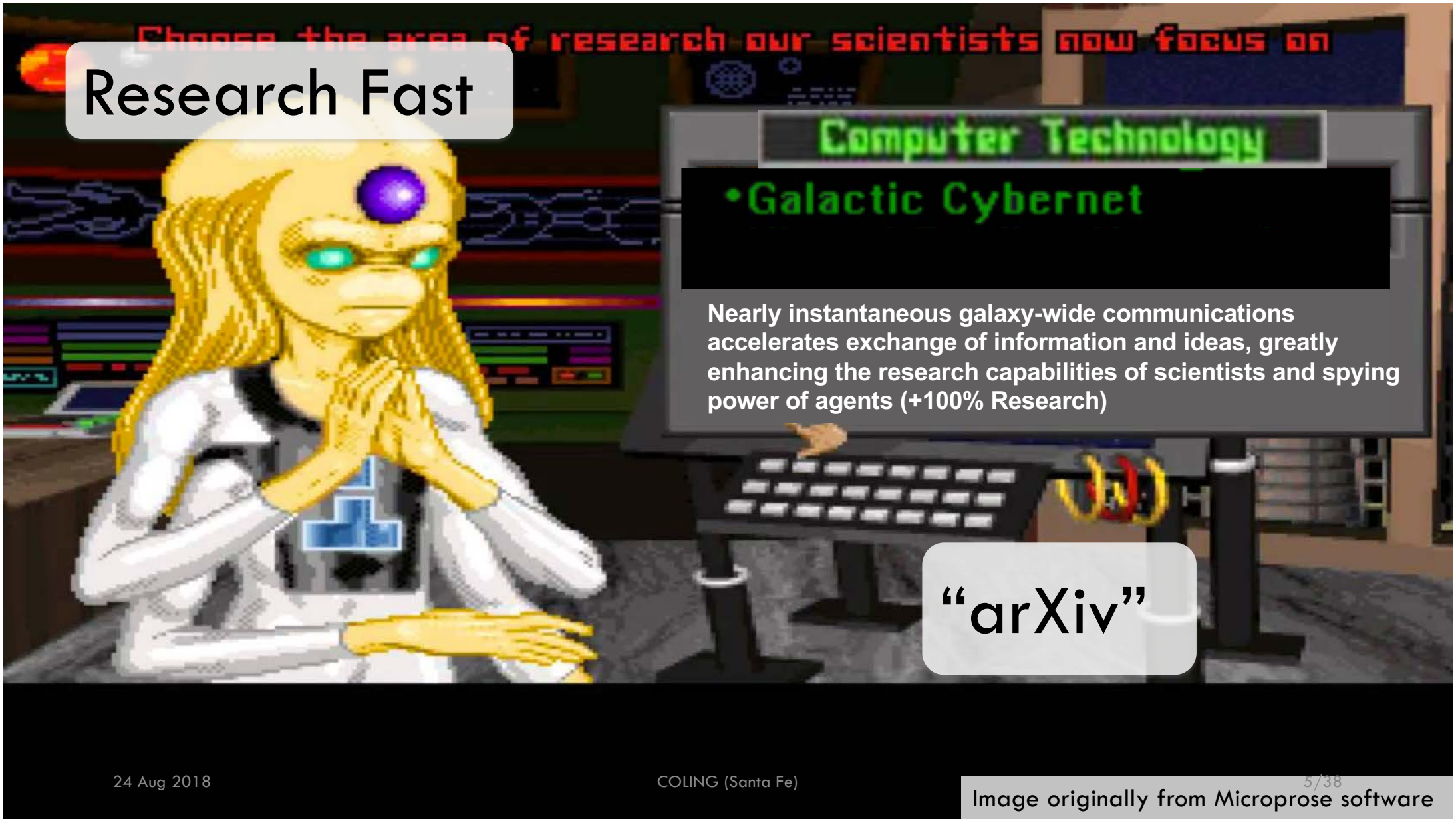
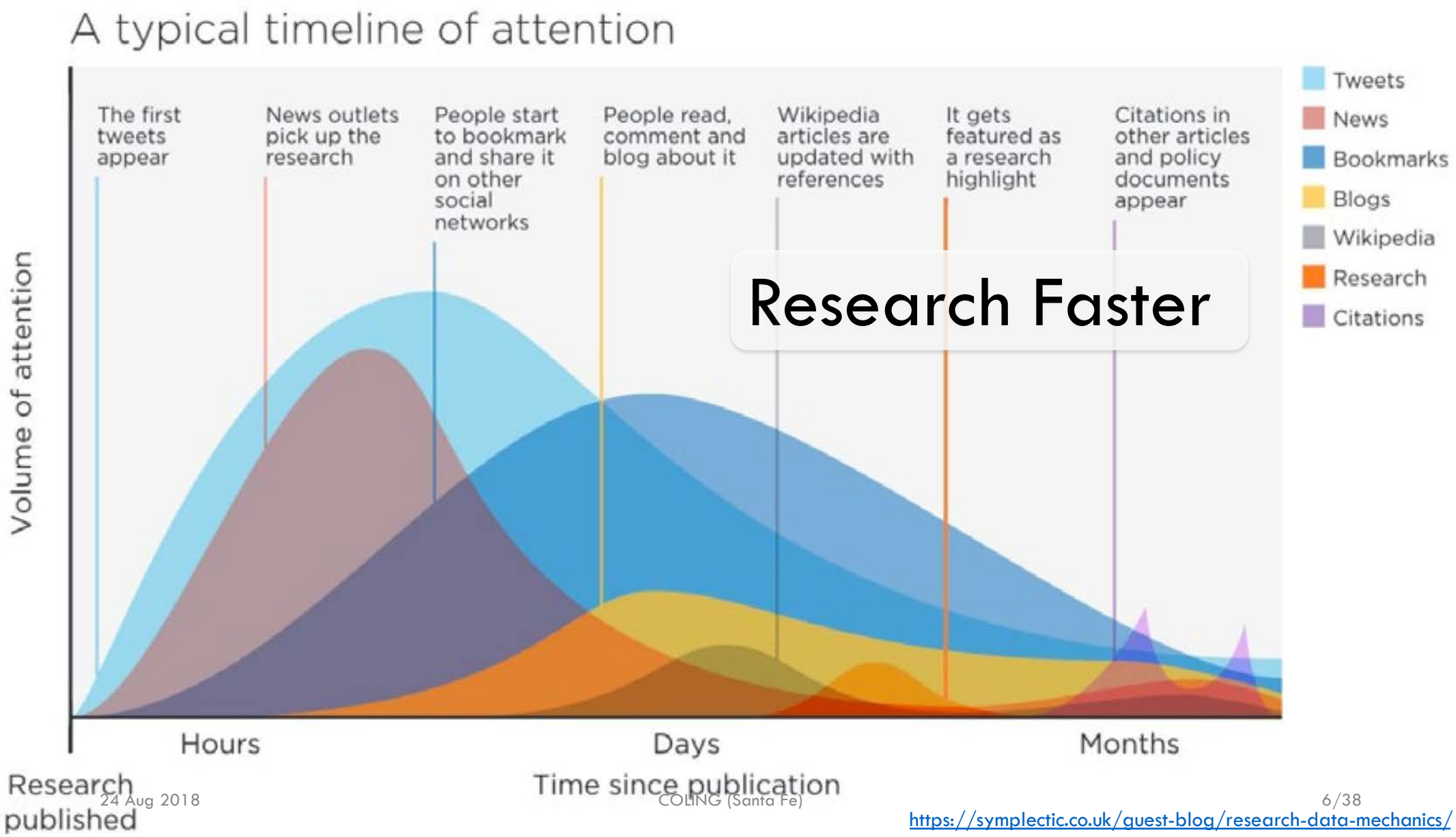


Image originally from Microprose software



Research Faster

https://symplectic.co.uk/guest-blog/research-data-mechanics/

Research Too Fast?

We have our own age of accelerations for research

1. arXiv – 驚輸 “kiasu”

2. GitHub / Jupyter Notebook

3. Shared Tasks

    Overheard Yejin Choi: “*Working on sQUaD makes people feel better*”

“*We shape our buildings, thereafter our buildings shapes us.*” – Winston Churchill

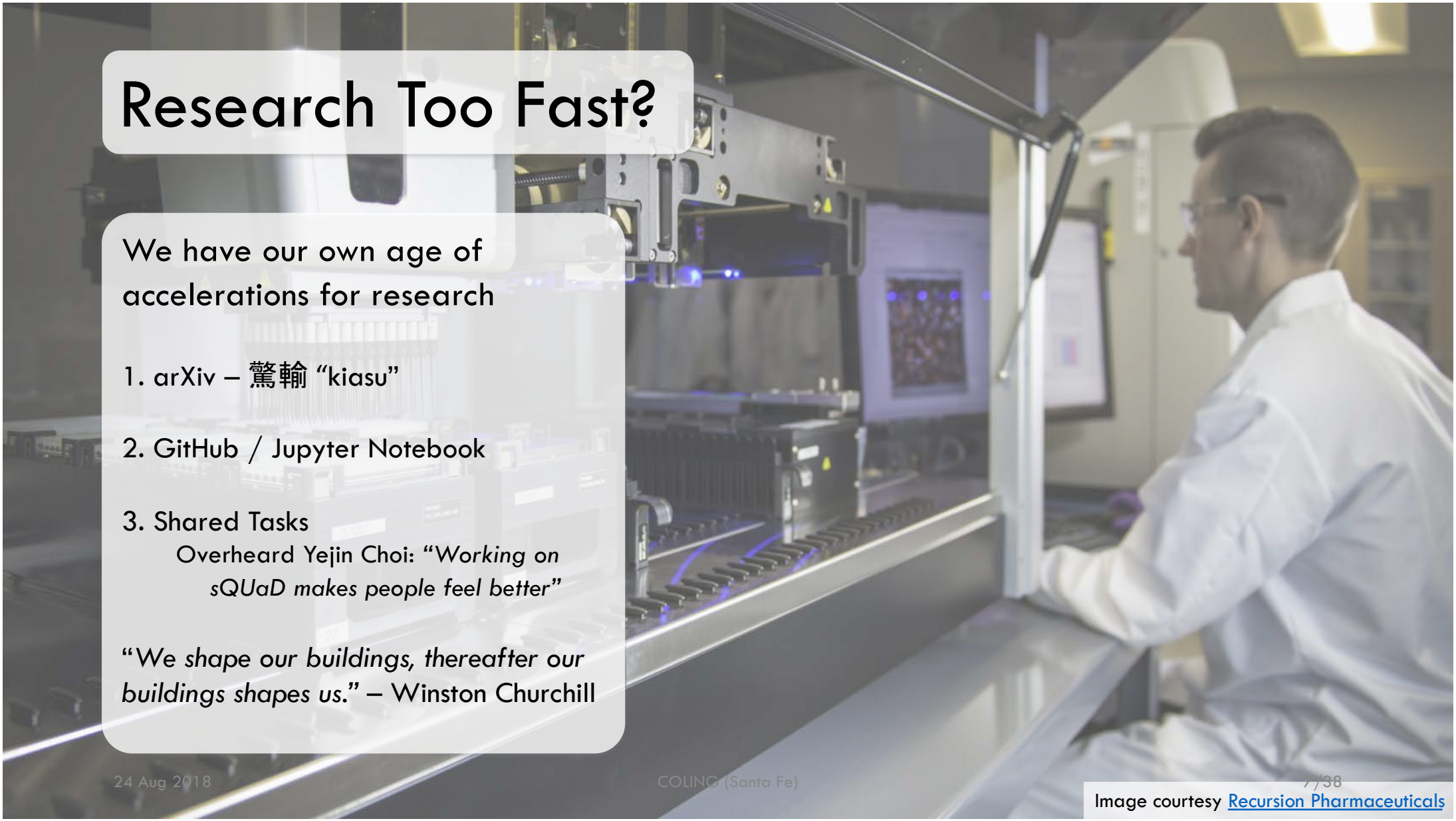
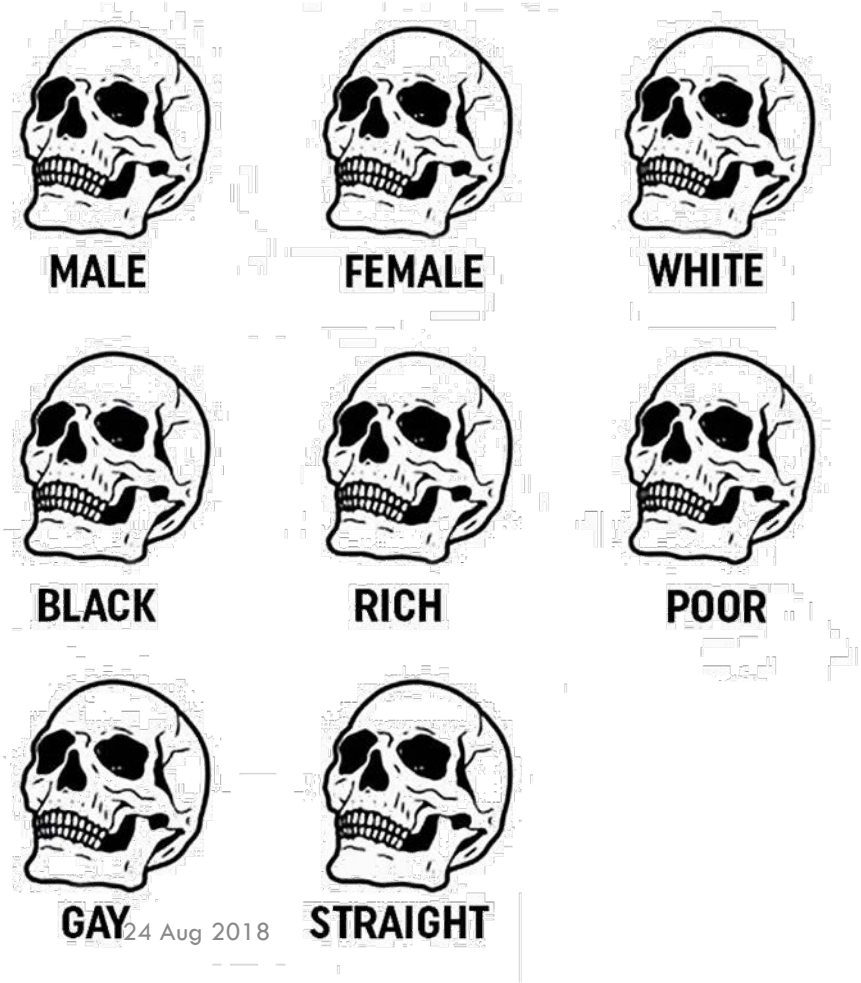


Image courtesy Recursion Pharmaceuticals

Here is a neat summary of the current state of interpretations of skull comparisons in biological anthropology:



Reviewer 2 must be stopped

Reviewers also favor defensible positions.

Safer to use standard metrics for acceptance.

Understandable for journals but for conferences?

Provenance unknown. Source: Reviewer 2 Must Be Stopped Facebook Group

Loss function of research

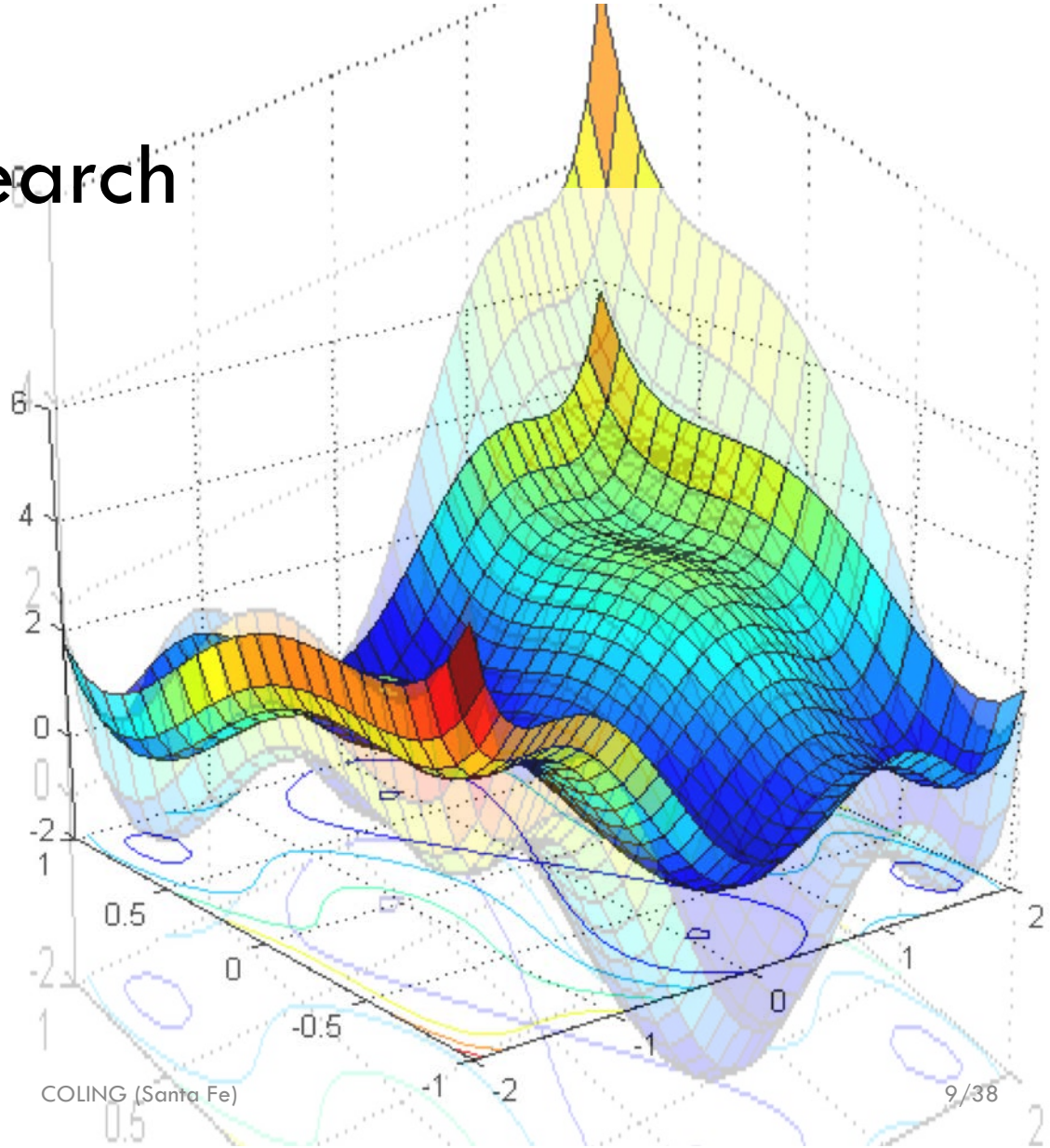
Beam search analogy

Accelerations make the gradient steeper

Overload favors the convenient

But scholars can't do random restart

Suboptimal local minima



Q: When your ML doesn't outperform, what do you do?

Your turn: select an answer.

1. Sigh.
2. Gather more data. Maybe your model needs more data to find proper weights.
3. Simplify your model. Maybe its easier just to do more hyperparameter searching.
4. Read your arXiv feed.

5. Study your problem more. Let your text speak.

StackExchange | Search on Linguistics… | 121 • 1 • 4

Linguistics beta

Questions | Tags | Users | Badges | Unanswered | **Ask Question**

What are the fundamental differences between Natural Language Processing and Computational Linguistics?

12

The above answers are all good. I'd like to offer another perspective that I learned while teaching digital libraries that draws on the analogy used in biology:

> Computational biology = the study of biology using computational techniques. The goal is to learn new biology, knowledge about living sytems. It is about science.
>
> Bioinformatics = the creation of tools (algorithms, databases) that solve problems. The goal is to build useful tools that work on biological data. It is about engineering.

To make the analogy for any field X, we thus have "Computational X" and "X-omatics". In NLP/CL, NLP is the equivalent of "Linguamatics".

I don't really subscribe to the notion that CL encompasses NLP or vice versa. They both have a purpose. CL studies human language to computationally understand how we as humans have the capacity to produce and understand language. NLP takes a more pragmatic perspective and says that we wish to build systems that facilitate some language interface.

share edit delete flag

answered Oct 27 '12 at 14:42

Min-Yen Kan

121 • 1 • 4

CL = Why?

NLP = What?

*Jason Eisner has a better answer on Quora

Let your text speak…but ask the right questions

How would a NLP system help scholars with this?

Digital library systems today are still monolithic scalable CMS

- They serve authors: prestige, directed discovery (search)
- And provision appropriate access and provenance: OpenURL, DOI

What about the users? We can do better.

Let's turn our science loose on our science.

Poster credit: JH Miller of Westinghouse Electric via Wikipedia

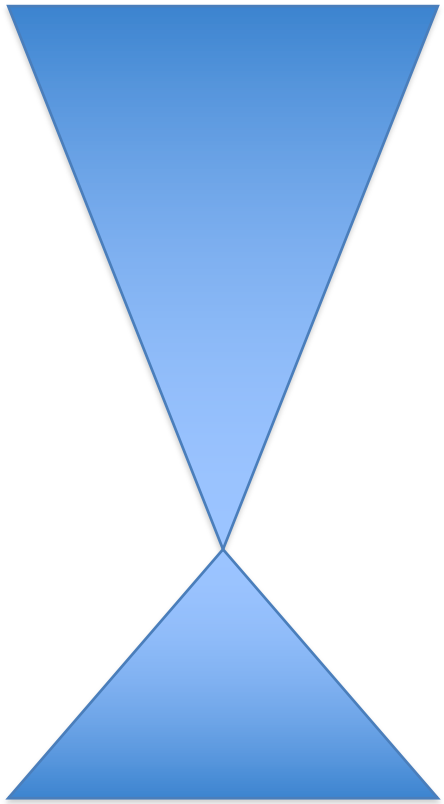
Towards the introspective

Digital Library

Photo Credits: sizumaru @ Flickr

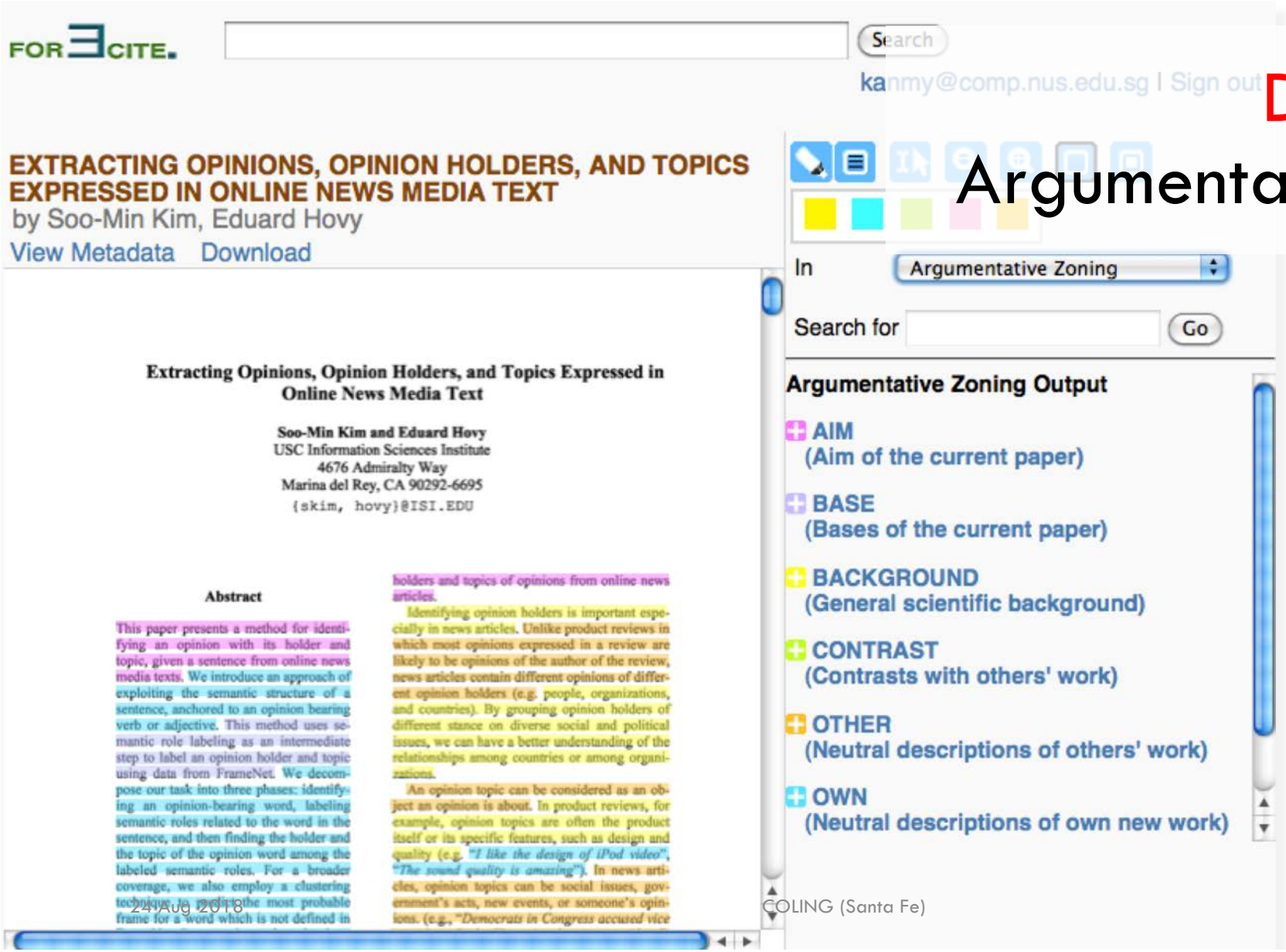
The (Slow) Process of Research

Enthusiasm
Discovering
Reading
Collecting Data
Sensemaking
Comparing
Publishing
Communicating
Maintaining

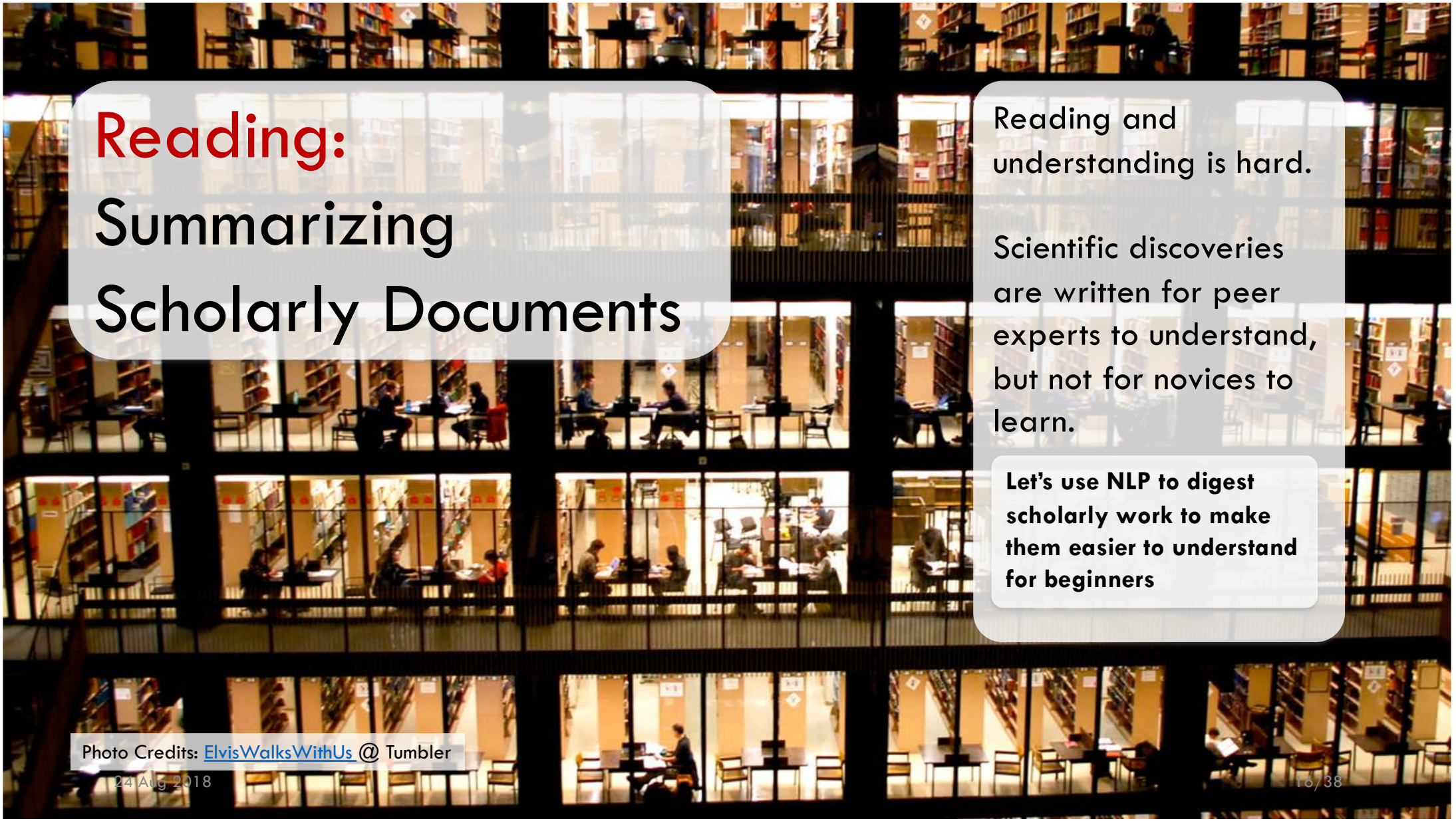


Title
Authors
Abstract
Introduction
Related Work
Method
Evaluation
References
Ancillary Artifacts

Discovering: Argumentative Zoning



Teufel and Kan (2011) Robust Argumentative Zoning for Sensemaking in Scholarly Documents.



Reading: Summarizing Scholarly Documents

Reading and understanding is hard.

Scientific discoveries are written for peer experts to understand, but not for novices to learn.

Let's use NLP to digest scholarly work to make them easier to understand for beginners

Photo Credits: ElvisWalksWithUs @ Tumbler

Summarizing Scholarly Documents

Why not just use the abstract?

The *bǎ* construction is central to the study of Mandarin grammar. It has received many attempts at analysis and comes up frequently as a syntactic test in discussions of other phenomena. Yet, not even its part of speech has ever been convincingly established. This paper presents the case for treating *bǎ* as a verb, considering both language-internal arguments and arguments from universal properties of parts of speech. These arguments are intended to have cross-theoretic validity. On the basis of the conclusion that *bǎ* is a verb, an analysis is developed within the framework of Lexical Functional Grammar. On this analysis, *bǎ* selects for a subject, an object, and a complement clause, and further stipulates that its object controls the TOPIC function of its complement clause. This analysis is shown to account for both the core data and the data which is problematic for other analyses. Finally, the analysis is confirmed by evidence from telicity effects in the *bǎ* construction, universal properties of verbs and prepositions, and its compatibility with the known historical development of the construction.

Rhetorical Structure Theory*

EMILY BENDER

THE SYNTAX OF MANDARIN *BǍ*:
RECONSIDERING THE VERBAL ANALYSIS*

The *bǎ* construction is central to the study of Mandarin grammar. It has received many attempts at analysis and comes up frequently as a syntactic test in discussions of other phenomena. Yet, not even its part of speech has ever been convincingly established. This paper presents the case for treating *bǎ* as a verb, considering both language-internal arguments and arguments from universal properties of parts of speech. These arguments are intended to have cross-theoretic validity. On the basis of the conclusion that *bǎ* is a verb, an analysis is developed within the framework of Lexical Functional Grammar. On this analysis, *bǎ* selects for a subject, an object, and a complement clause, and further stipulates that its object controls the TOPIC function of its complement clause. This analysis is shown to account for both the core data and the data which is problematic for other analyses. Finally, the analysis is confirmed by evidence from telicity effects in the *bǎ* construction, universal properties of verbs and prepositions, and its compatibility with the known historical development of the construction.

1. INTRODUCTION

The *bǎ* construction (1) is perhaps the most widely discussed phenomenon in modern studies of Mandarin Chinese. Nonetheless, there is little argument as to its proper analysis. *Bǎ* has been treated as a case marker for direct objects (Liang, 1971; Goodall, 1986; J. Huang, 1992); as a preposition (J. Huang, 1998; L.-Y. Huang, 1990; A. Li, 1990; and McCawley, 1992); as a "coverb" (Li and Thompson, 1981); as a verb (Hashimoto, 1971; Ross, 1991; Yang, 1995); and as a functional category heading its own projection (Zou, 1993; Sybesma, 1999).

This paper will argue for an analysis of *bǎ* as a verb and against the other possibilities. I will first bring together, refine, and add to the strongest of the existing arguments for the hypothesis that *bǎ* is a verb, and then provide an analysis based on the conclusions within the framework of Lexical-Functional Grammar (LFG) (Dalrymple et al., 1995; Bresnan, forthcoming).

The structure of the paper is as follows: Section 2 gives a brief sketch of the basic properties of the *bǎ* construction. Sections 3 through 5 present cross-theoretically valid arguments establishing the verbal status of *bǎ*. Section 6 presents some further evidence about the NPs which follow *bǎ*. Section 7 presents an analysis of the *bǎ* construction in the framework of LFG. This analysis will be shown to uniformly capture both "core" and

Journal of East Asian Linguistics 9, 105–145, 2000.

Let's look at our example

It follows a specific presentation convention

There is a logical document structure

INTRODUCTION

…

THE BASICS OF THE *BǍ* CONSTRUCTION

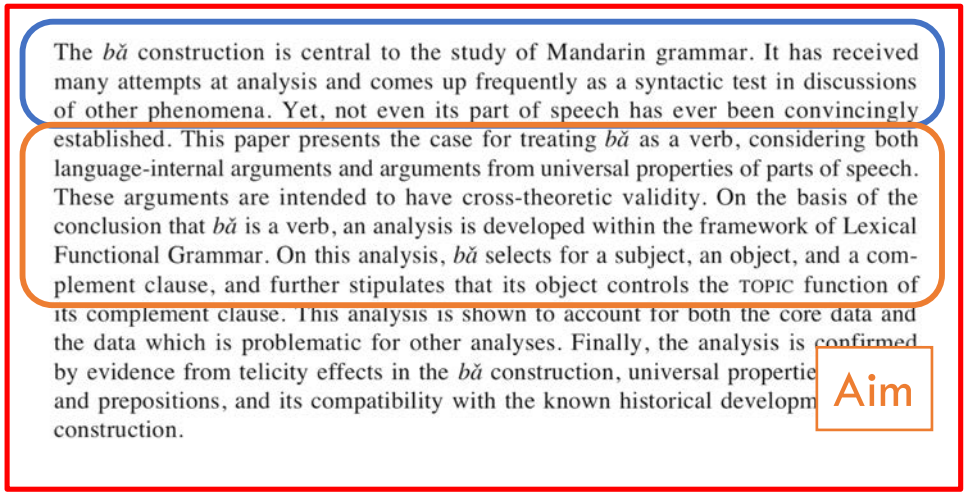
…

ANALYSIS

* *Apologies to Mann and Thompson*

Abstracting abstracts

Background

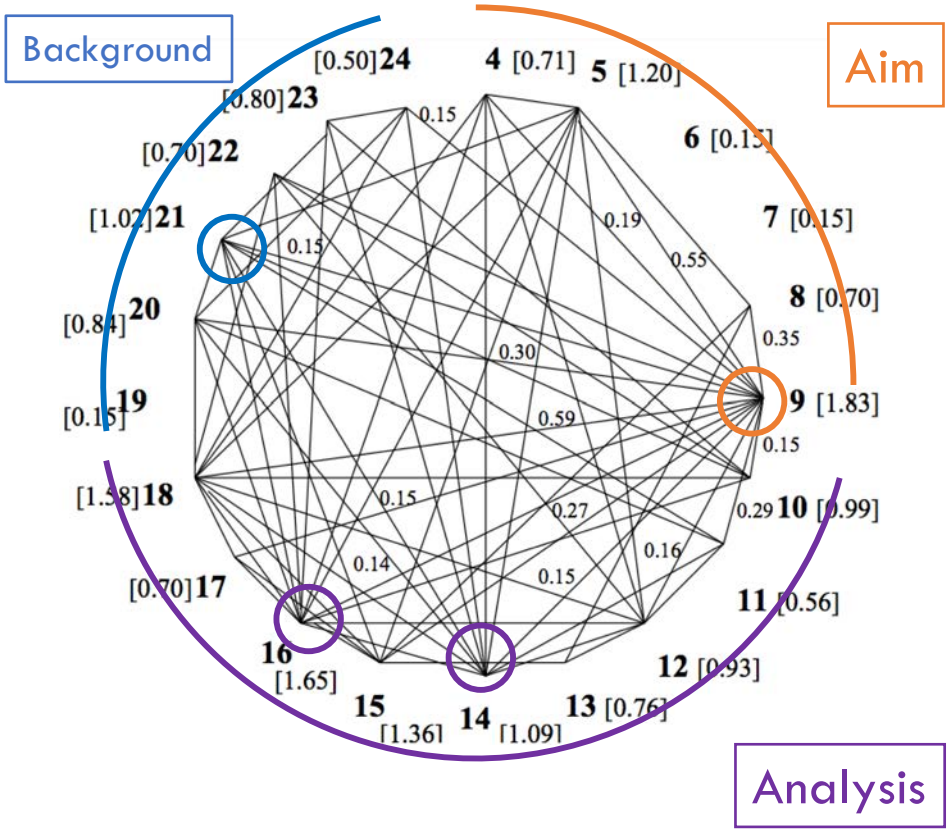


These are the "ground truth" for a summary of a single paper

Want to select sentences that overlap significantly with the abstract

But even here there is structure within an abstract

Revisiting Text Rank



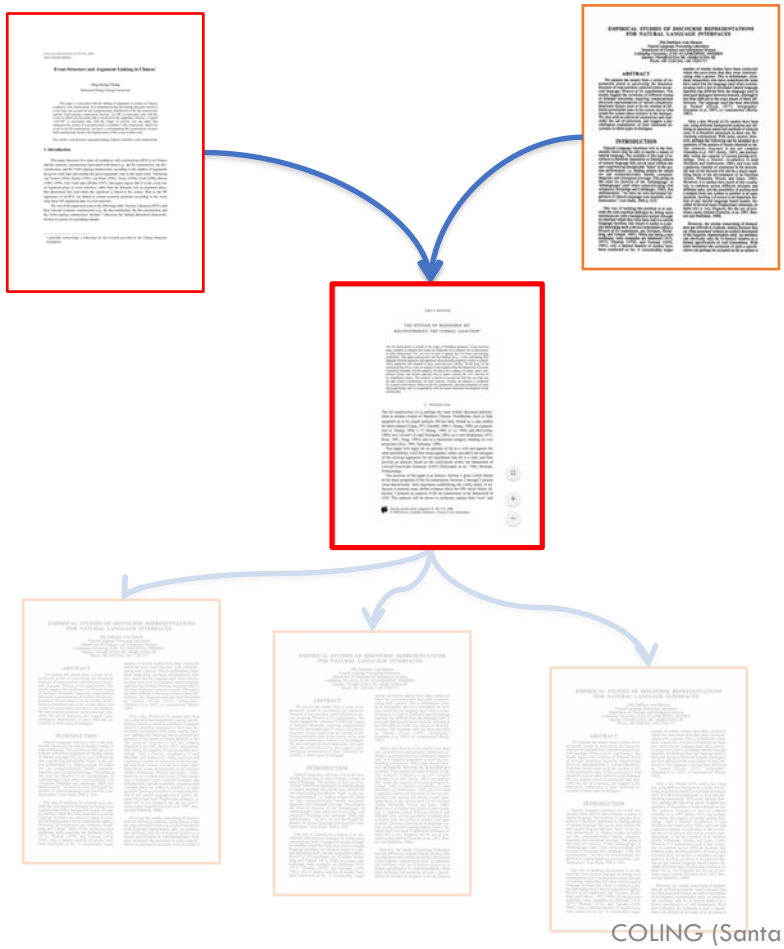
Capitalize on the conventional structure in documents

Identify the logical structure of a scientific document

Find the best sentences within the sections of the document

Jun Ping Ng, Praveen Bysani, Ziheng Lin, Min-Yen Kan and Chew Lim Tan (2011) SWING: Exploiting Category-Specific Information for Guided Summarization. In Proceedings of the Text Analysis Conference 2011 (TAC 2011). Gaithersburg, Maryland, USA.

What about *scholarly* documents?



They have references and citations "**citation sentences**"

- Often describe a paper from the community's point of view
- A representation of a key point of a work

Citation sentences and in- article sentences have complementary purposes:

- Results and evaluations usually not mentioned in citation sentences.
- Sentences in a paper that describe its method are usually too detailed for a summary.

Emily M. Bender

FOLLOW

Professor, Dept of Linguistics, University of Washington
Verified email at uw.edu - Homepage

Computational Linguistics Linguistics

TITLE	CITED BY	YEAR
Syntactic theory: A formal introduction IA Sag, T Wasow, EM Bender, IA Sag Center for the Study of Language and Information	1582	1999
The grammar matrix: An open-source starter-kit for the rapid development of cross-linguistically consistent broad-coverage precision grammars EM Bender, D Flickinger, S Oepen Proceedings of the 2002 workshop on Grammar engineering and evaluation …	275	2002
Efficient deep processing of Japanese M Siegel, EM Bender Proceedings of the 3rd workshop on Asian language resources and …	142	2002
Syntactic variation and linguistic competence: The case of AAVE copula absence EM Bender stanford university	116	2000
The syntax of Mandarin Bă: Reconsidering the verbal analysis E Bender Journal of East Asian Linguistics 9 (2), 105-145	90	2000
Grammar customization EM Bender, S Drellishak, A Fokkens, L Poulson, S Saleem Research on Language and Computation 8 (1), 23-72	75	2010
Rapid prototyping of scalable grammars: Towards modularity in extensions to a language-independent core EM Bender, D Flickinger Companion Volume to the Proceedings of Conference including Posters/Demos …	74	2005
Arboretum: Using a precision grammar for grammar checking in CALL EM Bender, D Flickinger, S Oepen, A Walsh, T Baldwin Instil/icall symposium 2004	68	2004
Road-testing the English Resource Grammar Over the British National Corpus. T Baldwin, EM Bender, D Flickinger, A Kim, S Oepen LREC	67	2004

Sensemaking: Citation Analysis

What do they refer to when they cite?
- Citation Provenance.
The whole paper? A specific section of the paper?

Why do people cite?
- Citation Function.
What functions are there?

lead to false independence assumptions. For example, the BA marker is descriptively used in Chinese to preverbalize objects. Its syntax, however, is controversial (Bender, 2001). In the

In addition, we proposed a set of new features that used verb class information induced from the frame files of the Chinese PropBank, as well as features that were designed to exploit the grammatical constructions that Semantic Role Labeling of Chinese Predicates are unique to Chinese, specifically the BA (Bender 2000) and BEI (Huang 1999) constructions.

For discussion of the modern Mandarin disposal construction the reader is referred to Li and Thompson (1981), Cheng (1988), Sybesma (1999), Bender (2000), among many others.

In the literature, the word *ba* has been treated in many different ways. For example, it is treated as a case marker for direct objects (Liang 1971, Goodall 1986, J. Huang 1992, Y.Y. Huang 1991), as a preposition (J. Huang 1982, L-Y. Huang 1990, Y.-H. Li 1990, Li and Thompson 1976, McCawley 1992), as a coverb (Li and Thompson 1981), as a verb (Hashimoto 1971, Ross 1991, Yang 1995, Bender 2000), as a secondary topic marker (Tsao 1987b), and as a functional category heading its own projection (Zou 1993, Sybesma 1999). In my analysis, the word *ba* is used to mark the displaced NP

The syntax of Mandarin Bă: Reconsidering the verbal analysis
E Bender
Journal of East Asian Linguistics 9 (2), 105-145

The context of the citation influences its importance: via its function and referent

- Its location (which section)
- Its syntactic structure

Low, H. W. (2011). Citation provenance. Bachelor's thesis, National University of Singapore.
Yulianto, E. (2012). Citation typing. Bachelor's thesis, National University of Singapore.

Discovery: Serendipitous Recommendation

Interactions with others play an important role in populating the slow hunch

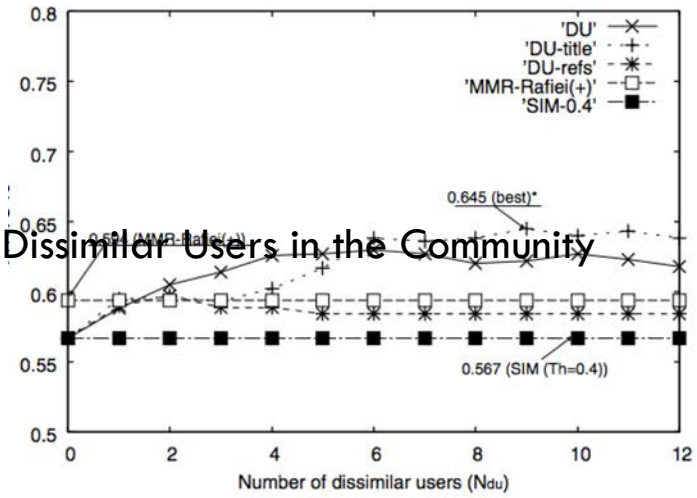
Photo Credits: Virtueel Platform @ Flickr, CC BY-SA 2.0

Kazunari Sugiyama and Min-Yen Kan (2011) Serendipitous Recommendation for Scholarly Papers Considering Relations Among Researchers. In Proceedings of the 11th ACM/IEEE Joint Conference on Digital Libraries (JCDL 2011) Short Papers.

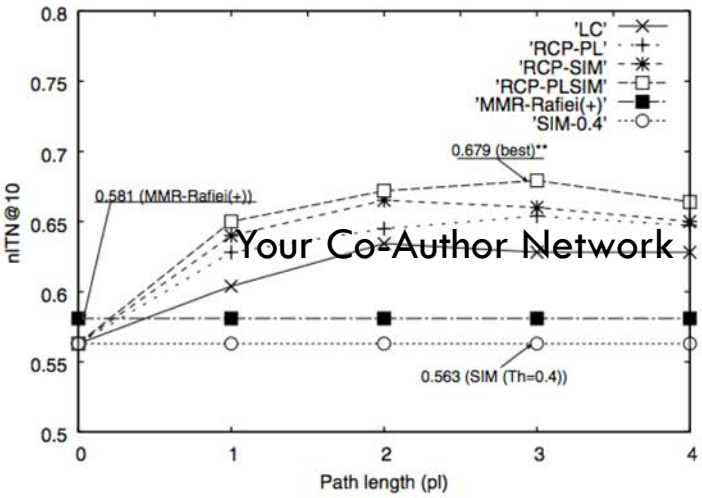
Tell me something I don't know

Collaborative Filtering: "People like you also read…"

Q: Whom do we ask for help?



Dissimilar Users in the Community



Your Co-Author Network

Sensemaking: Conference Presentations

We attend conferences (like this one) in part to help learn from each other.

A key artifact is the slide presentation, which often summarizes the work in an accessible manner.

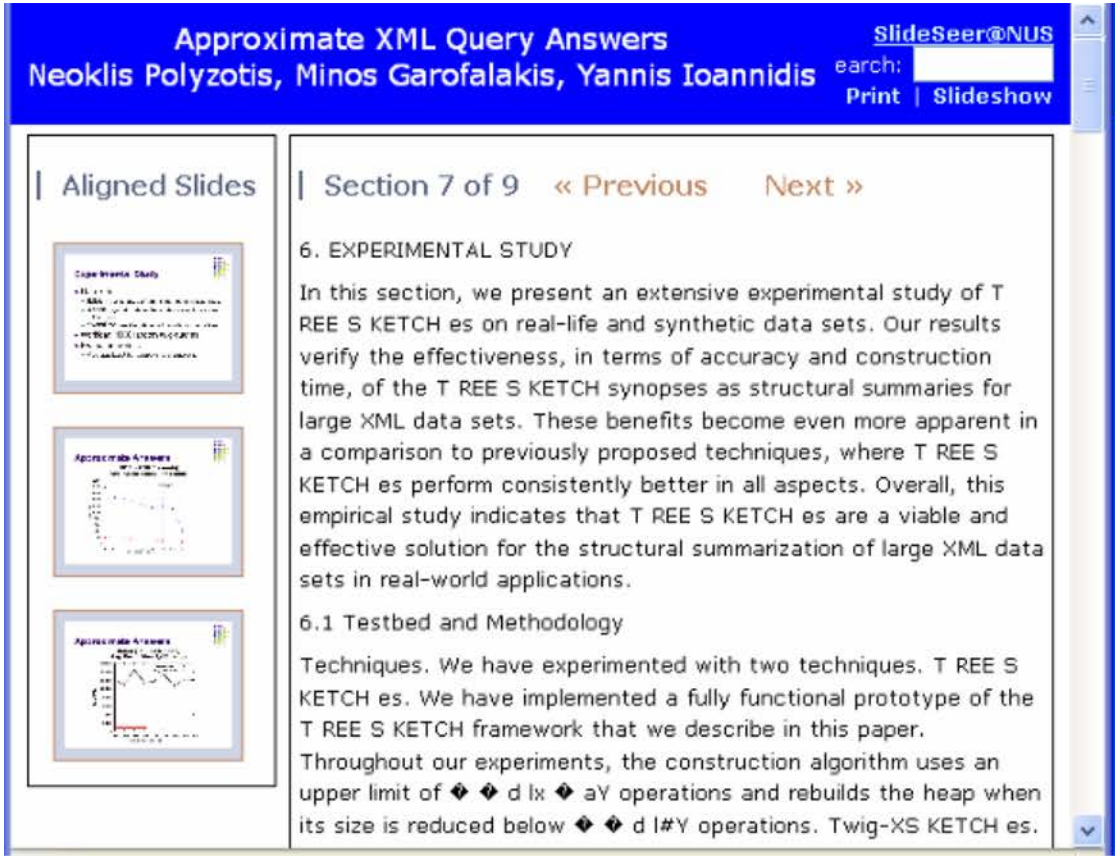
But they:

- Miss important technical details

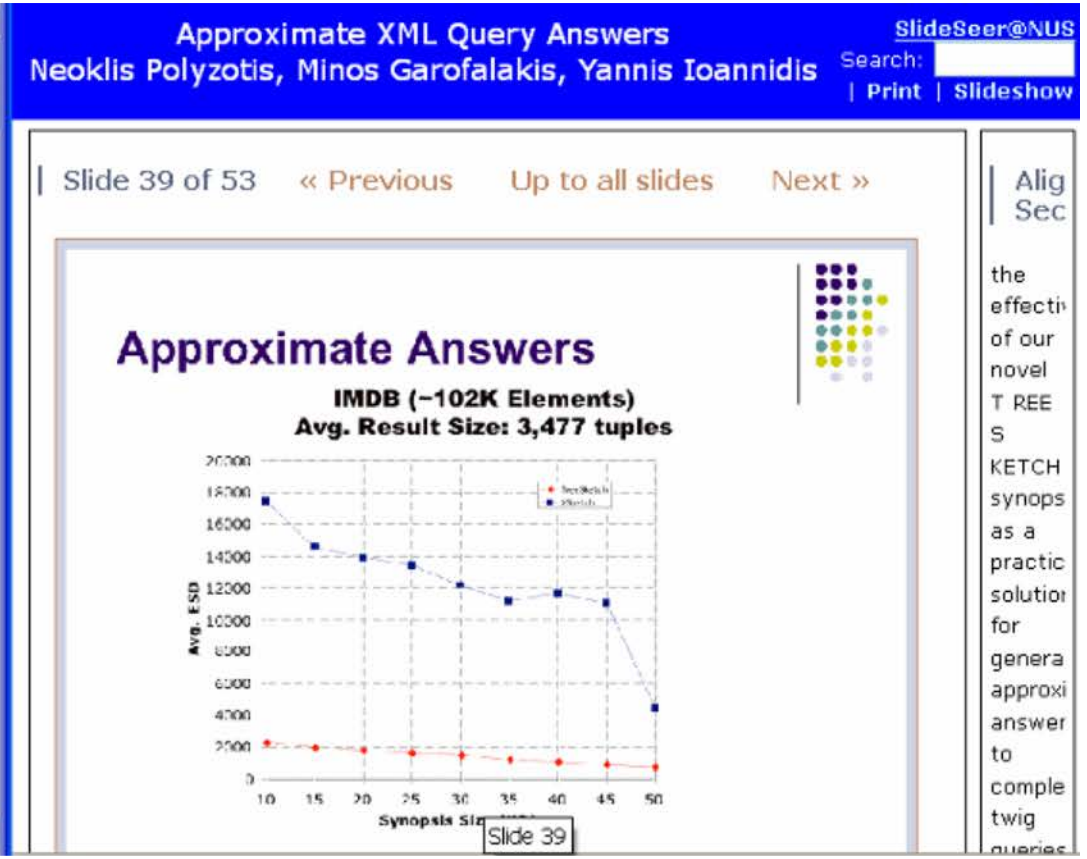
Idea: Use both together

Photo by myself, cleared to use by subject

Presentations and their Relationship to Documents



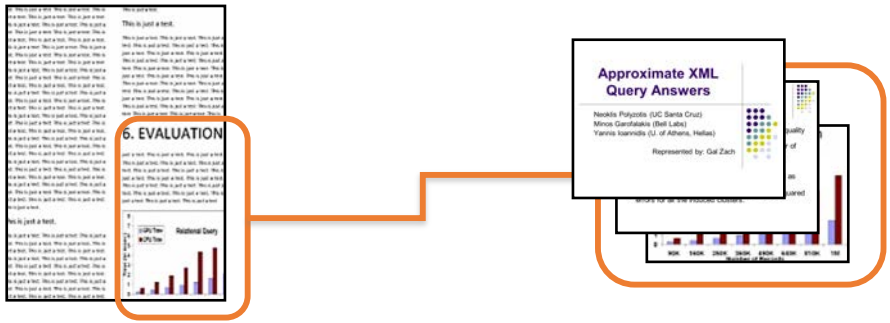
Document in focus



Presentation in Focus

Aligning Documents to their Presentations

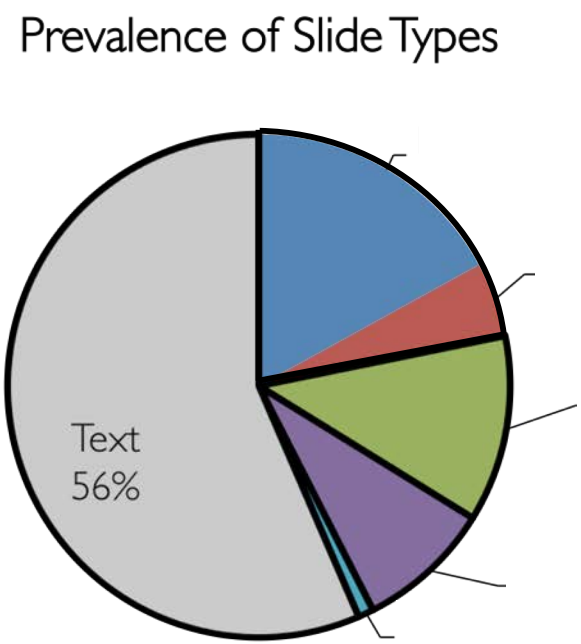
Better to juxtapose both media together in a fine-grained manner.

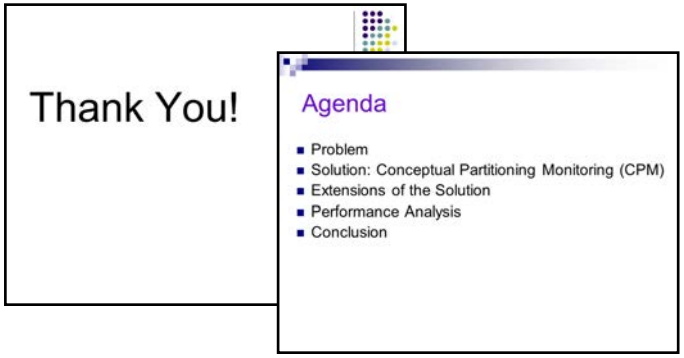


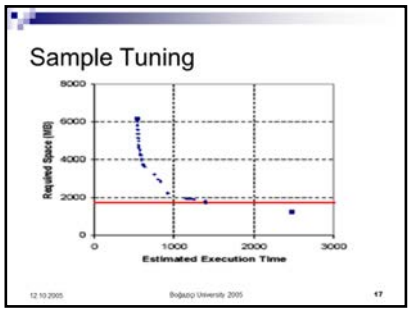
Output: an alignment map

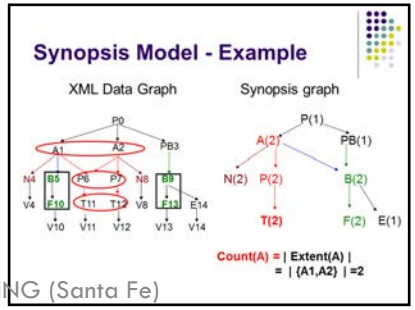
Bamdad Bahrani and Min-Yen Kan (2013) Multimodal Alignment of Scholarly Documents and Their Presentations. In Proceedings of the Joint Conference on Digital Libraries (JCDL '13). 22-26 July, Indianapolis, USA. pp. 281-284. Short Paper.
Min-Yen Kan (2007) SlideSeer: A Digital Library of Aligned Document and Presentation Pairs, In Proceedings of the Joint Conference on Digital Libraries (JCDL '07). Vancouver, Canada, June, pp. 81-90.

Slide Demographics









Error Analysis: It's a multimodal problem

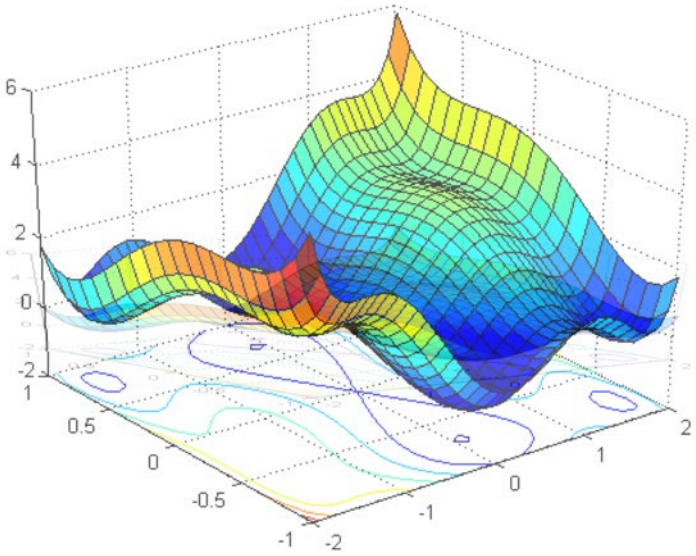
Slide Type	Common reason	% Incorrectly Aligned by Baseline
Nil	Doesn't know where to align → align to best fit	64%
Outline	Name of some sections in it → align to longest one	36%
Image	Very little text available	**81%**
Drawing	Noisy data: lots of shapes and text boxes	**53%**
Table	Little text, noisy data	50%
Text		24%

Research Slow

Ignore the gradient:
Open up the Adjacent Possible

Develops the Slow Hunch through persistent exposure to ancillary evidence

The Swedish Fika: collision allows scholars to benefit from each other



Where Good Ideas Come From

Steven Johnson

- The Adjacent Possible
- Liquid Networks
- The Slow Hunch
- Serendipity
- Error
- Exaptation
- Platforms

Fuben-Eki
不便益

FUrther **BENE**fit of
a **K**ind of **I**nconvenience

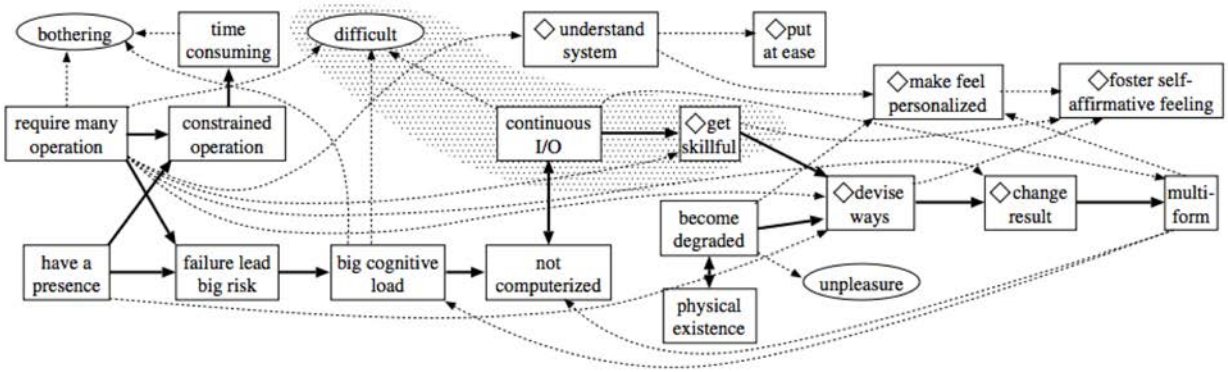


Fig. 1. Relations among inconvenience and benefits

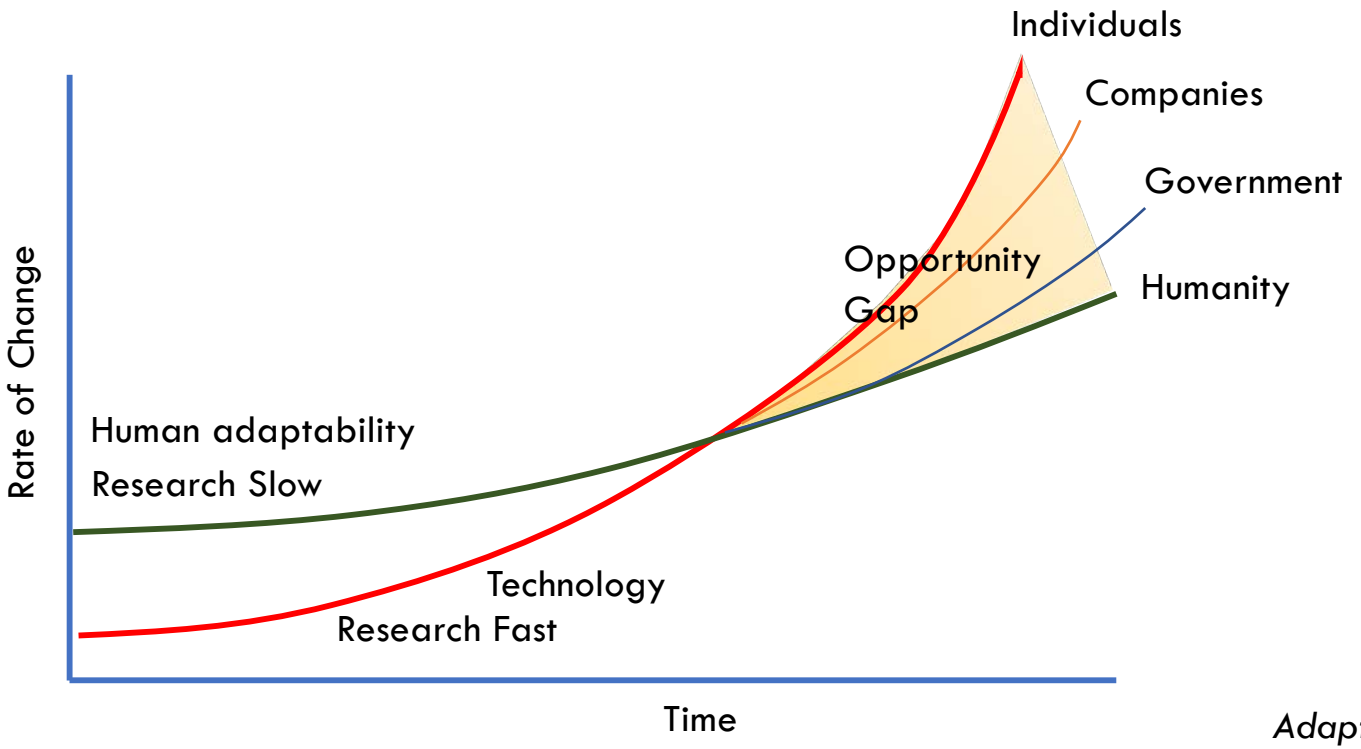
Make key operations deliberate

Fosters ownership, self-affirmation

Key idea: synergize with human computational abilities

Kawakami H. Toward System Design based on Benefit of Inconvenience, Journal of Human Interface, 11(1), 125-134 (2009) (in Japanese).

Acceleration has it out for us



Others are public advocates, but what about ourselves?

We need to ramp up our research slow

Use the time gifted by research fast

Adapted from Astro Teller's graph

Weapons of Math Destruction

Cathy O'Neil

A WMD is

- Massive
- Opaque
- No feedback loop

The Class Break

Reprise: Research Slow – The Challenge of Story

Be an advocate

- Write the story, wonder “what if?”
- Blog your opinion or write columns

Stories give context

- We are human
- Aids association, catalyzes the slow hunch

Ensembles are better than any single classifier

Hiyao Miyazaki: *Castle in the Sky*
宮崎 駿　　天空の城ラピュタ

Slides: http://bit.ly/kan-coling18

Conclusion: We combine theory and practice

… nothing works and we don't know why
… v2: It works but we don't know why
… v3: It works and we think we know why

… v4: It works and we think we know why and we'll advocate for it

Let's fix that: Research Fast AND Slow

Fast begets fast, but we need to consciously support slow

NLP to do intelligence augmentation (IA), put scholars on the slow hunch

Close the loop: Let data speak and inform our models

Thanks to:

WING members:
Bamdad Bahrani
Muthu Chandrasekaran
Tao Chen
Emily Bender
Leon Derczynski
Pierre Isabelle
Marti Hearst
Gertjan van Noord
Chris Manning
Xiangnan He
Aminesh Prasad
Kazunari Sugiyama
Aobo Wang
Yohei Seki

all those students and faculty at NUS whom I subjected to a dry run

And research slow technology of audiobooks and digital commonplace books