Computer Science Research 101

Zixian Ma | University of Washington

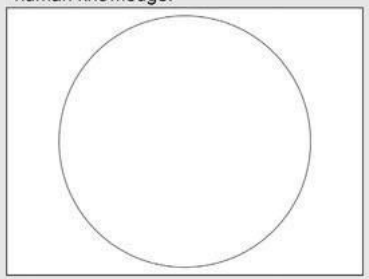
Adapted from Prof. Michael Bernstein's slides for CS 197 at Stanford University

Outline

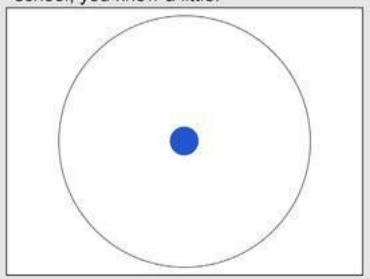
- Why (CS) research
- What is CS research
- Research contributions: Bit flips
- Literature search: Related work
- Explaining a project: Introduction
- Designing an Evaluation

Why (CS) research?

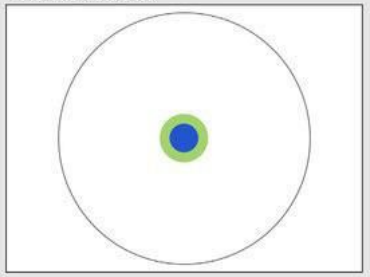
Imagine a circle that contains all of human knowledge:



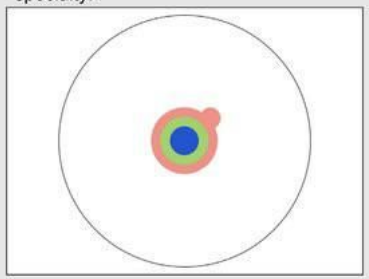
By the time you finish elementary school, you know a little:



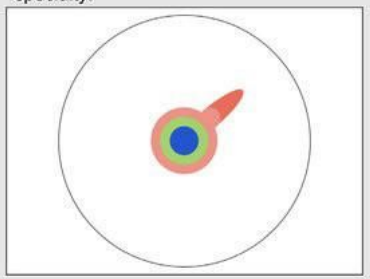
By the time you finish high school, you know a bit more:



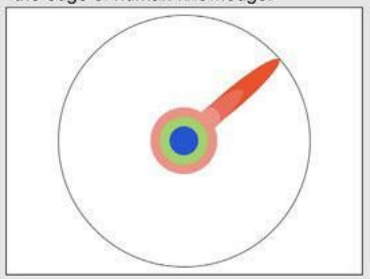
With a bachelor's degree, you gain a specialty:



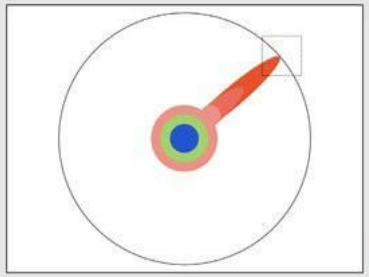
A master's degree deepens that specialty:



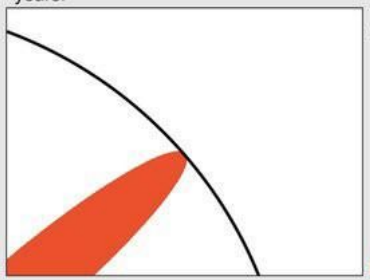
Reading research papers takes you to the edge of human knowledge:



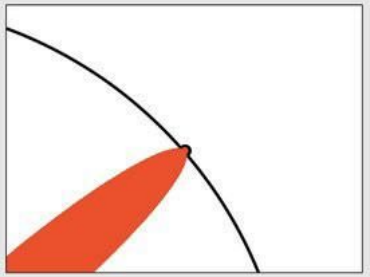
Once you're at the boundary, you focus:



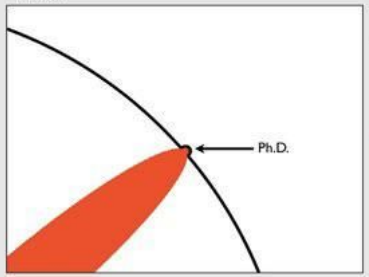
You push at the boundary for several years:



Until one day, the boundary gives way:

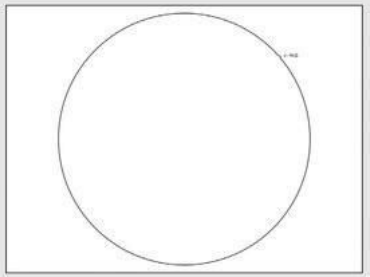


And that dent you've made is called a Ph.D.



Of course, the world looks different to you now:

So, don't forget the bigger picture:



What is CS research?

Seeking a Better Way to Find Web Images

By JOHN MARKOFF NOV. 19, 2012

STANFORD, Calif. — You may think you can find almost anything on the Internet.

But even as images and video rapidly come to dominate the Web, search engines can ordinarily find a given image only if the text entered by a searcher matches the text with which it was labeled. And the labels can be unreliable, unhelpful ("fuzzy" instead of "rabbit") or simply nonexistent.

To eliminate those limits, scientists will need to create a new generation of visual search technologies — or else, as the Stanford computer scientist Fei-Fei Li recently put it, the Web will be in danger of "going dark."

Now, along with computer scientists from Princeton, Dr. Li, 36, has bu
the world's largest visual database in an effort to mimic the human vis
system. With more than 14 million labeled objects, from obsidian to
orangutans to ocelots, the database has because a vital resource for
computer vision researchers.

CLOUD COMPUTING

Making Cloud-Computing Systems More Efficient

BY QUENTIN HARDY MARCH 6, 2014 7:00 AM 3

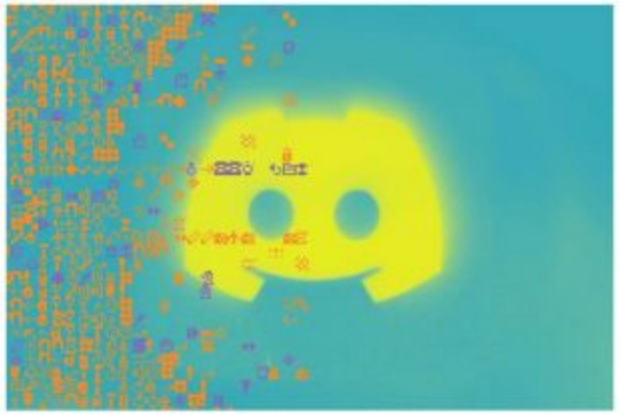
Email
Share
Tweet
Save
More

Christos Kozyrakis, professor of electrical engineering and computer science at Stanford University, headed the creation of management software called Quasar.

Computer scientists at Stanford have developed software that works like the recommendation engines for Amazon shoppers and Netflix movie subscribers, only for the distribution of workloads across large computing environments. The results, they say, can triple the efficiency of cloud-computing systems.

The New York Times

Plus: How A.I. is bringing us closer to "Westworld."



Kirill Kudryavtsev/Agence France-Presse — Getty Images

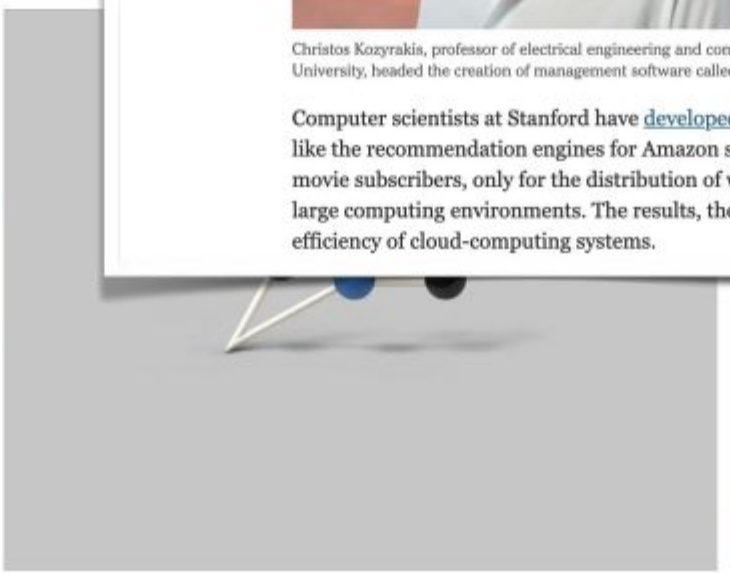
Stanford Researcher Finds Lots of Leaky Web Sites

BY SOMINI SENGUPTA OCTOBER 11, 2011 6:32 PM 8

The Web is porous. Remarkable information trickles in from everywhere. It also sometimes spills out without its users knowir exactly where or how.

Take for instance these findings, released on Tuesday by comput scientists at Stanford University. If you type a wrong password ir the Web site of The Wall Street Journal, it turns out that your e-mail address quietly slips out to seven unrelated Web sites. Sign to NBC and, likewise, seven other companies can capture your e-mail address. Click on an ad on HomeDepot.com and your first name and user ID are instantly revealed to 13 other companies.

These findings, released by the Center for Internet and Society at Stanford Law School, are among the leaks found on 185 top Web



Doug Chayka

A.I. Is Learning What It Means to Be Alive

Given troves of data about genes and cells, A.I. models have made some surprising discoveries. What could they

The New York Times

SCIENCETAKE

OceanOne, a Mer-Bot Dive Buddy With a 'Friendly Face'

Computer science research

What is the goal of research?

Why has it driven major innovations in computing?

What separates research from advanced development?

A Tale of Two Turing Awards

Engelbart: interactive computing

When computers originated, they were used for, well, computing: calculating mathematical functions.

This meant that computers were seen as most appropriate for slow, batch interaction, shared by entire teams.

The New York Times

GIVE THE TIMES

DOUGLAS C. ENGELBART, 1925-2013

Computer Visionary Who Invented the Mouse

By John Markoff

July 3, 2013

Douglas C. Engelbart was 25, just engaged to be married and thinking about his future when he had an epiphany in 1950 that would change the world.

He had a good job working at a government aerospace laboratory in California, but he wanted to do something more with his life, something of value that might last, even outlive him. Then it came to him. In a single stroke he had what might be safely called a complete vision of the information age.

The epiphany spoke to him of technology's potential to expand human intelligence, and from it he spun out a career that indeed had lasting impact. It led to a host of inventions that became the basis for the Internet and the modern personal computer.

In later years, one of those inventions was given a warmhearted name, evoking a small, furry creature

Engelbart: interactive computing

"No, let's do it this way instead:" computing should be used as a tool for thought. We must move from batch-style computing to interactive computing.

His result was the "Mother of All Demos": mouse, hypertext, bitmapped screens, collaborative software, and more.

This led to Xerox Star. Steve Jobs saw it, was wow'ed, and infused the ideas into the Macintosh.

The New York Times

GIVE THE TIMES

DOUGLAS C. ENGELBART, 1925-2013

Computer Visionary Who Invented the Mouse

By John Markoff

July 3, 2013

Douglas C. Engelbart was 25, just engaged to be married and thinking about his future when he had an epiphany in 1950 that would change the world.

He had a good job working at a government aerospace laboratory in California, but he wanted to do something more with his life, something of value that might last, even outlive him. Then it came to him. In a single stroke he had what might be safely called a complete vision of the information age.

The epiphany spoke to him of technology's potential to expand human intelligence, and from it he spun out a career that indeed had lasting impact. It led to a host of inventions that became the basis for the Internet and the modern personal computer.

In later years, one of those inventions was given a warmhearted name, evoking a small, furry creature

"The Mother of All Demos"

LeCun, Hinton, Bengio: deep learning

The idea of neural networks had been around for fifty years, but unsuccessful. Major AI figures had trashed it, even proving that early versions had very limited expressiveness.

Instead, machine learning was based on other models, for example the support vector machine and graphical models. Neural networks did not perform well.

The New York Times

GIVE THE TIMES

Turing Award Won by 3 Pioneers in Artificial Intelligence

From left, Yann LeCun, Geoffrey Hinton and Yoshua Bengio. The researchers worked on key developments for neural networks, which are reshaping how computer systems are built.

From left, Facebook, via Associated Press; Aaron Vincent Elkaim for The New York Times; Chad Buchanan/Getty Images

By Cade Metz

March 27, 2019

LeCun, Hinton, Bengio: deep learning

"No, let's do it this way instead:" these networks learn extremely complex functions, so they need much more data than existing machine learning approaches, GPUs to train, and algorithms to enable them to learn more effectively.

Around 2010, these models began smashing records in speech and image recognition. They are now foundational to ML.

The New York Times

GIVE THE TIMES

Turing Award Won by 3 Pioneers in Artificial Intelligence

From left, Yann LeCun, Geoffrey Hinton and Yoshua Bengio. The researchers worked on key developments for neural networks, which are reshaping how computer systems are built.

From left, Facebook, via Associated Press; Aaron Vincent Elkaim for The New York Times; Chad Buchanan/Getty Images

By Cade Metz

March 27, 2019

Not all research wins Turing Awards. But…

It all follows this same formula —

An implicit assumption: Industry and other researchers all thought one way about a problem

“No, let's do it this way instead:” The researcher offered a new perspective that nobody had ever considered or made feasible before. They proved out their idea as the better approach.

And now, a definition.

Research introduces a fundamental new idea into the world.

Examples:

- Computing that is interactive, not batch
- Algorithms needed to make deep learning effective

These ideas did not exist in any mature or well-articulated way before their creators developed them.

If the idea is already in the world, for example published by someone else, it is not considered novel, and thus not research.

Before: small computer vision datasets

After: huge computer vision dataset, and algorithms to utilize it

Before: programmers manually reserve resources for cloud computing

After: programmers provide needs, software allocates

Before: models of people are sparse

After: generative AI agents capture broad behavior

Before: we think web tracking is isolated to the intended site

After: it's much leakier than we realized

Before: biologists had to discover cell functions

After: AI can discover cell functions

Before: underwater robots should look and feel like boats

After: humanoid underwater robotics

Research creates industry

PageRank algorithm

Apache Spark: big data processing engine

Computer virtualization

Online education

Early "gateway" (router) software

Industry vs. research

What makes other start-ups and industry different than research?

If the core idea already exists, but needs to be refined in order to see success…it might be important, but it's not research.

(An incomplete list of) research areas in computer science

Architecture
Artificial intelligence
Computational biology
Computer graphics
Computer security
Computer systems
Computer vision
Data science
Education
Human-computer interaction
Machine learning
Natural language processing
Networking
Operating/distributed systems
Programming systems/verification
Robotics
Theory

Architecture

Artificial intelligence

Computational biology

Computer graphics

Computer security

Computer systems

Computer vision

Data science

Education

Human-computer interaction

Machine learning

Natural language processing

Networking

Operating/distributed systems

Programming systems/verification

Robotics

Theory

Topic: artificial intelligence

Architecture
Artificial intelligence
Computational biology
Computer graphics
Computer security
Computer systems
Computer vision
Data science
Education
Human-computer interaction
Machine learning
Natural language processing
Networking
Operating/distributed systems
Programming systems/verification
Robotics
Theory

Topic: computer systems

Architecture
Artificial intelligence
Computational biology
Computer graphics
Computer security
Computer systems
Computer vision
Data science
Education
Human-computer interaction
Machine learning
Natural language processing
Networking
Operating/distributed systems
Programming systems/verification
Robotics
Theory

Topic: theory

Architecture

Artificial intelligence

Computational biology

Computer graphics

Computer security

Computer systems

Computer vision

Data science

Education

Human-computer interaction

Machine learning

Natural language processing

Networking

Operating/distributed systems

Programming systems/verification

Robotics

Theory

Method: probability and modeling

Architecture
Artificial intelligence
Computational biology
Computer graphics
Computer security
Computer systems
Computer vision
Data science
Education
Human-computer interaction
Machine learning
Natural language processing
Networking
Operating/distributed systems
Programming systems/verification
Robotics
Theory

Method: design

Architecture
Artificial intelligence
Computational biology
Computer graphics
Computer security
Computer systems
Computer vision
Data science
Education
Human-computer interaction
Machine learning
Natural language processing
Networking
Operating/distributed systems
Programming systems/verification
Robotics
Theory

Method: empirical measurement and hypothesis testing

Research mindset

Research is different than your usual coursework.

- Coursework tends to be very clearly defined. Research tends to be exploratory and iterative.
- You probably won't know the right answer: if we knew whether it was going to work, it wouldn't be research.

In other words, there must be novelty

If the idea is already in the world, it is not considered novel, and thus not research.

To do research, you need to achieve something that nobody else has ever done. That novel achievement is called the contribution of your research.

You'll hear people say things like:

"This is an extremely novel contribution."

"This work is a tad too incremental." *(its improvement or level of creativity over the state of the art is only minimal)*

Novelty how? Flip The Bit

Bit flip: invert an assumption

An inversion of an assumption that the world has about how the world is supposed to work.

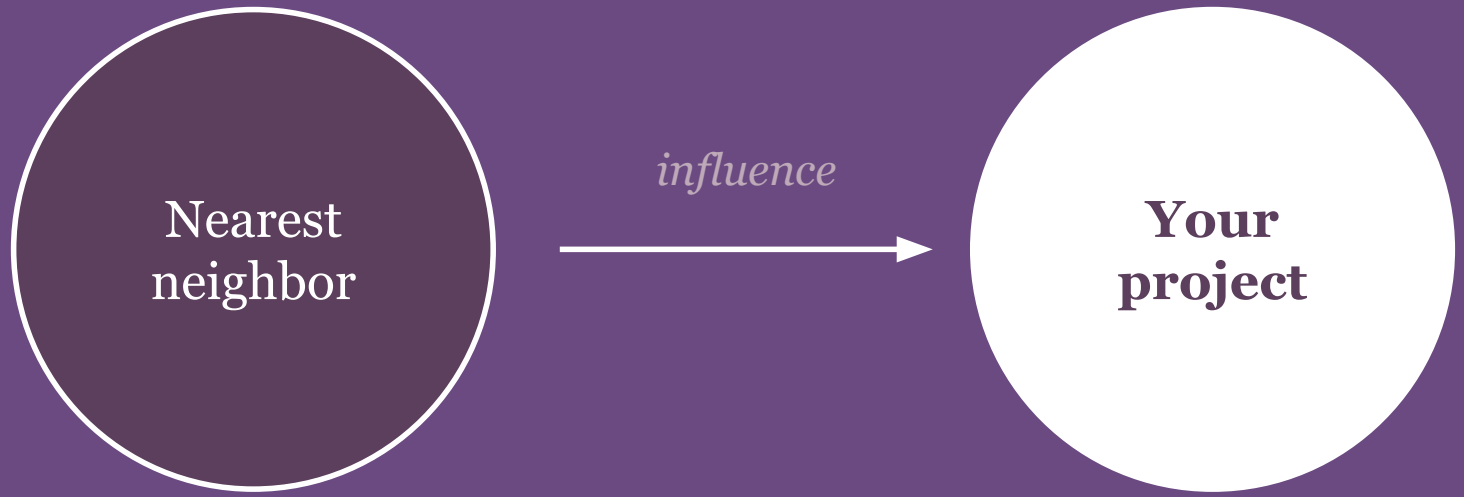
Recipe for a bit flip:

1) Articulate an assumption, often left implicit in prior work: this is the bit

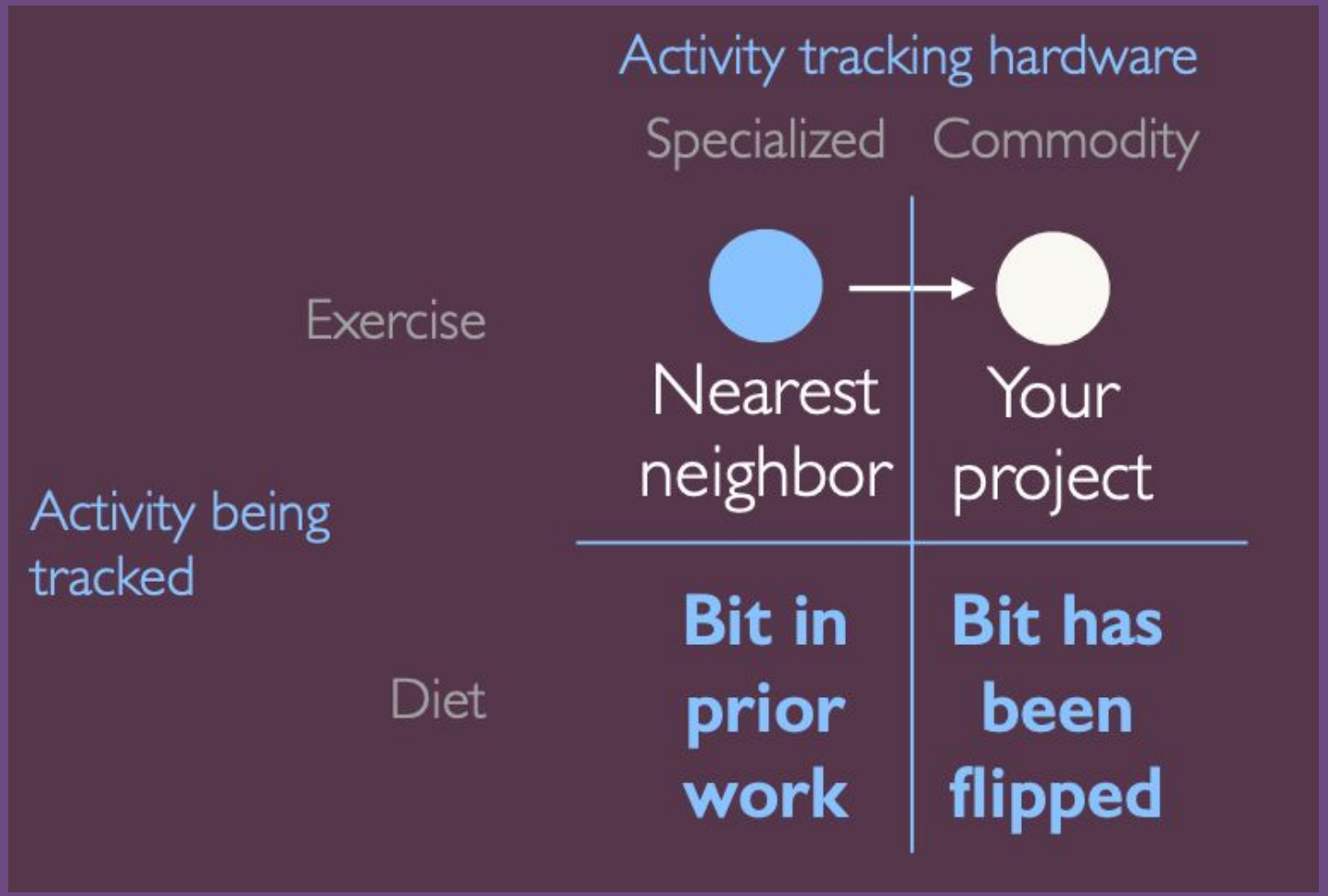
2) *"No, it should be this way instead:"* argue for an alternative to that assumption

Project	*Bit*	*Flip*
Ubifit	Activity tracking requires custom hardware.	Activity tracking requires just a standard cell phone.
BERT	NLP machine learning models should read sentences word by word, so the model can see what's before the current word.	NLP machine learning models should consume the entire sentence at once, so the parser can see what's before and after.

First, know the bit via literature search



Literature search graph



Imagine a set of design axes.

Your project should maintain position on most of them, but differentiate itself along one axis.

Single paper bit flip: exercise

Vision Transformer paper [Dosovitskiy et al. ’20]:

An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit, Neil Houlsby

While the Transformer architecture has become the de-facto standard for natural language processing tasks, its applications to computer vision remain limited. In vision, attention is either applied in conjunction with convolutional networks, or used to replace certain components of convolutional networks while keeping their overall structure in place. We show that this reliance on CNNs is not necessary and a pure transformer applied directly to sequences of image patches can perform very well on image classification tasks. When pre-trained on large amounts of data and transferred to multiple mid-sized or small image recognition benchmarks (ImageNet, CIFAR-100, VTAB, etc.), Vision Transformer (ViT) attains excellent results compared to state-of-the-art convolutional networks while requiring substantially fewer computational resources to train.

What are the bit and flip? [2 mins]

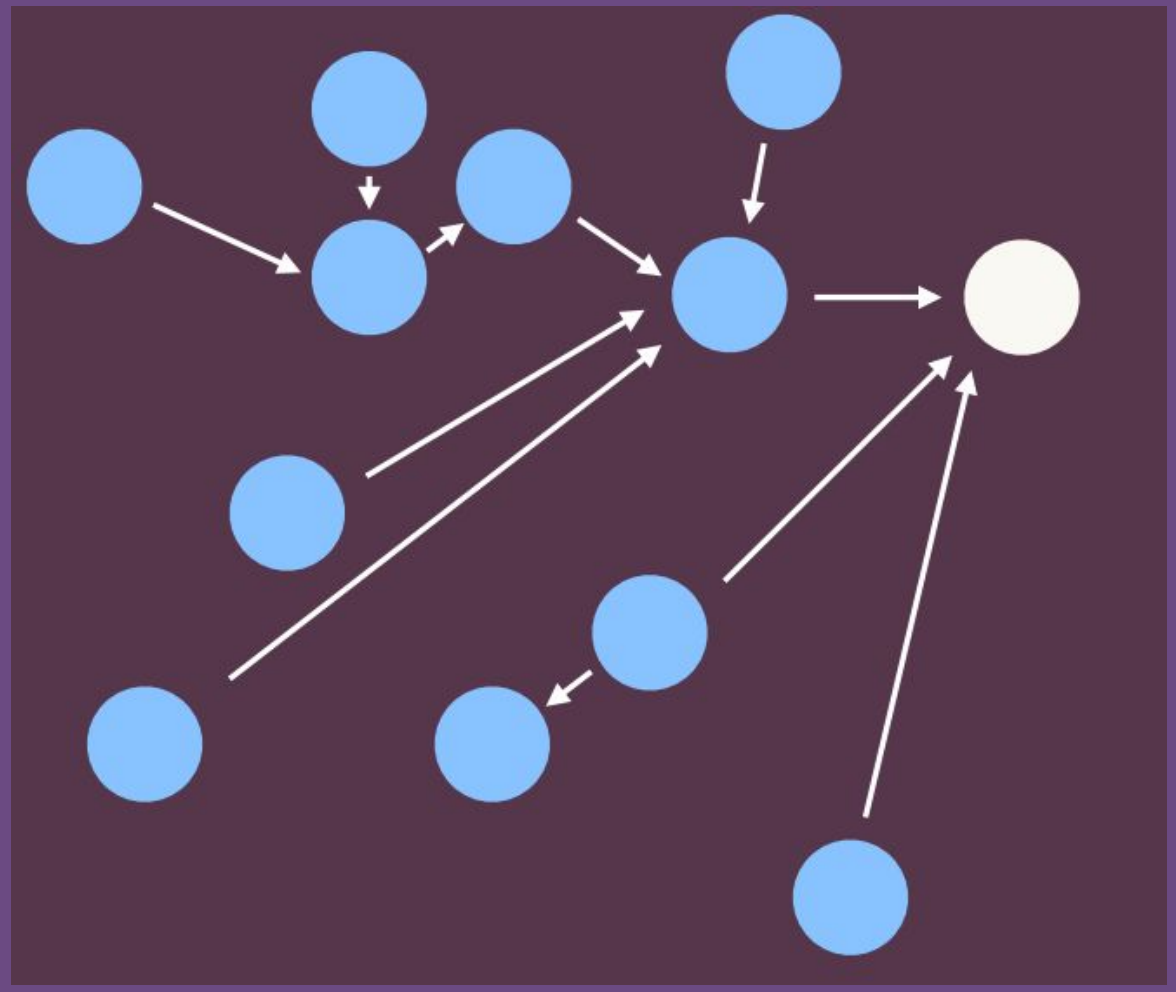
Literature-level bit flip

Eventually your goal is not to pivot off a single paper, but off the literature more broadly. This makes for a stronger argument of novelty.

Recipe:

1) Read the literature. (Which papers?)
2) What assumptions underlie all of the papers? $\forall\, p \in papers$
3) Which assumption are you changing? And why does it matter to the literature?

Literature: clearer bit flip



Why do a literature-level bit flip instead of a paper-level bit flip?

There exist many possible bit flips for a single paper, but not every possible bit flip matters. Some are incremental.

The broader an understanding you have of the literature and the design axes underneath it, the more effectively you can pick the right bit flip.

Ultimately…

It’s unlikely that you will find an idea that nobody has ever articulated in any context ever.

Instead, your goal is to articulate the broadest class of papers possible that your bit flip applies to.

Explaining a project → Introduction

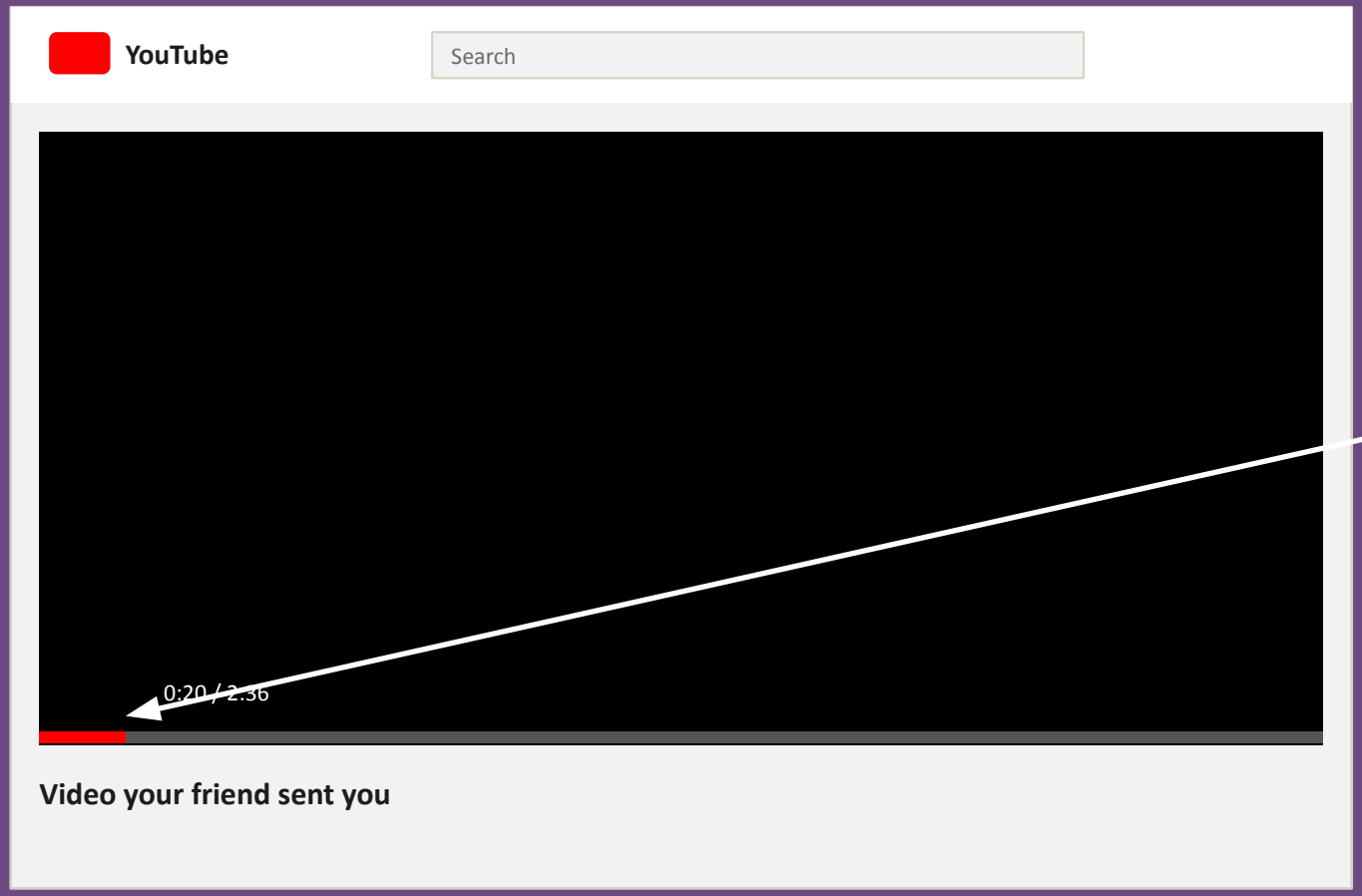
Try it!

Turn to a partner and explain your project to them [1 min each]

- How clearly do you understand the problem that your partner is addressing?
- How clearly do you understand the solution that your partner is proposing?

What is an Introduction?

- *The Introduction makes the case for your research, in brief.*
- One way to think about it:



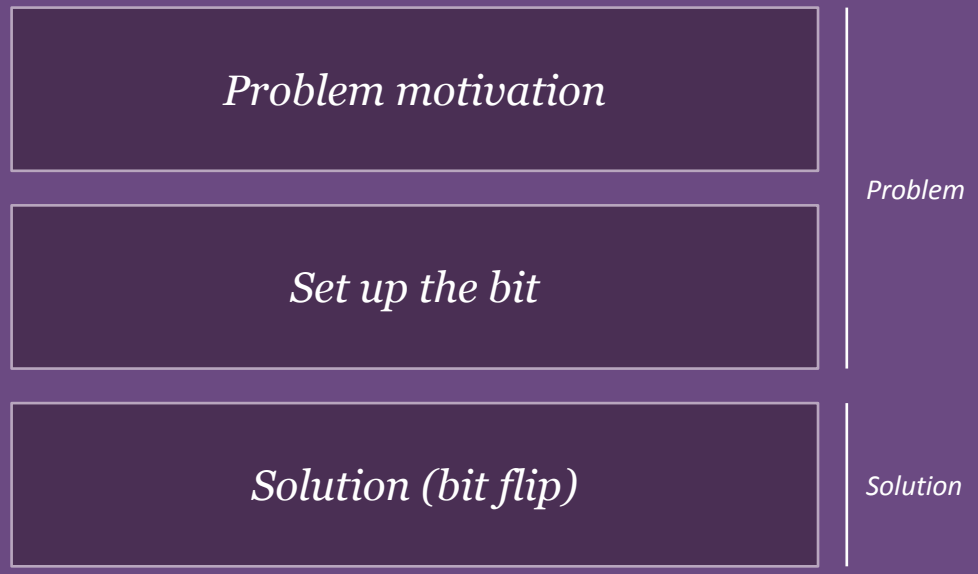
By this point, the video has hopefully made clear to you what it's about, and you've made a decision about whether to watch the rest of it.

Each introduction makes the case for two things:

- The problem: Why do we care about the problem you're solving?
- The solution: Why is your approach creative and correct?

Architecture of an intro

…great, Zixian thanks. But how do we actually do this?



Problem motivation

Explain the main problem that you're trying to solve.

Use citations to back up your claims about the existence of the problem, and why we should care about solving it.

Problem motivation

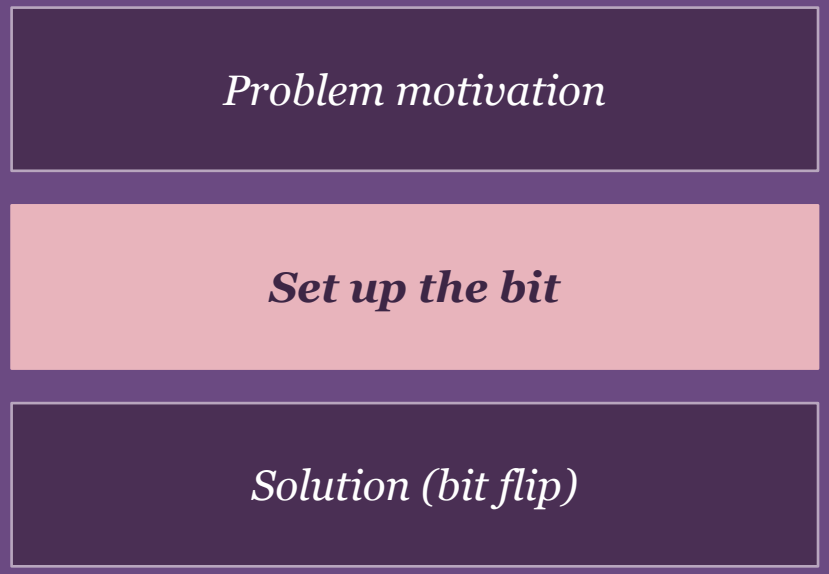
Set up the bit

Solution (bit flip)

Set up the bit

Answer the question, "Why isn't this problem solved yet?" by setting up the bit that you're going to flip. This is a summary of related work — one that's directly in service of your bit.

- *Networks are hard to (re)configure*
- *Interactions with computers are stuck on flat glass displays*
- *Generative AI models are challenging to evaluate*



Crowdsourcing platforms such as Amazon Mechanical Turk decentralize their workforce, designing for distributed, independent work [16, 42]. Decentralization aims to encourage accuracy through independent judgement [59]. However, by making communication and coordination more difficult, decentralization disempowers workers and forces worker collectives off-platform [41, 64, 16]. The result is disenfranchisement [22, 55] and an unfavorable workplace environment [41, 42]. Worse, while decentralization is motivated by a desire for high-quality work, it paradoxically undercuts behaviors and institutions that are critical to high-quality work. In many traditional organizations, for example, centralized worker coordination is a keystone to behaviors that improve work quality, including skill development [2], knowledge management [35], and performance ratings [58].

bit = decentralization

The rest of this introduction paragraph is dedicated to surveying related work with respect to how decentralization is architected, and to its outcomes.

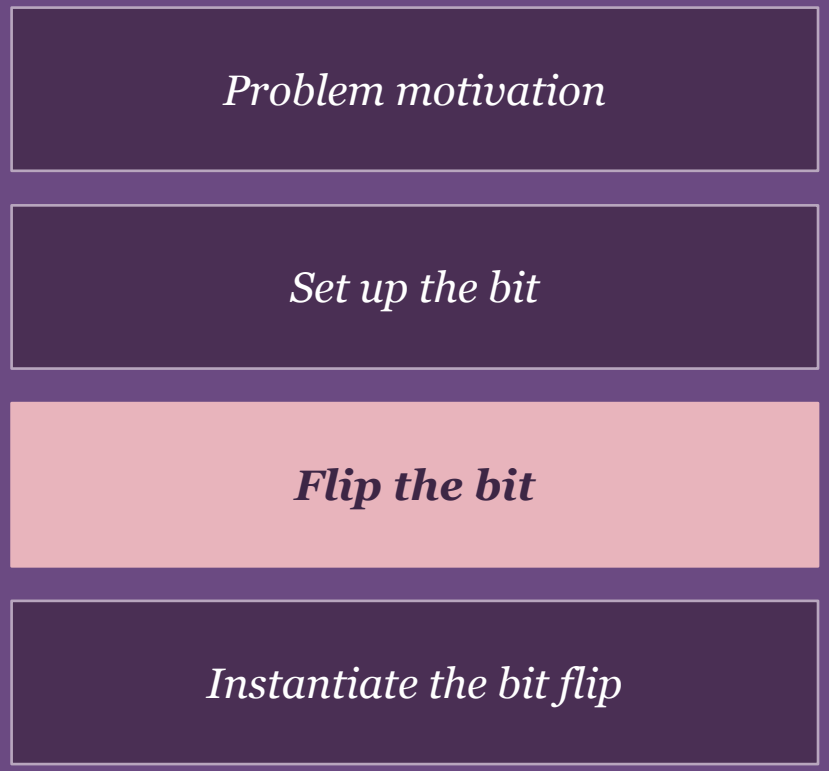
Unpacking the solution

The solution has to explain two things: what the big idea is, and how that big idea gets instantiated in the specific context of this problem.

(Even if someone hears your bit flip — that you want to introduce recurrence inside the neural network — they may still have no idea how that actually connects to the problem of language generation.)

Flip the bit

- The topic sentence of this paragraph is the thesis statement of your entire research project.
- Pivot off of the bit you set up to flip the bit. Explain why flipping the bit is a good idea for the problem at hand.
- It should now be obvious to a reader given the prior paragraph that this research is novel, since you have proven that nobody else has flipped that bit.



To address this reputation challenge, and with an eye toward other challenges that arise from decentralization, we draw inspiration from a historical labor strategy for coordinating a decentralized workforce: *guilds*. Worker guilds arose in the early Middle Ages, when workers in a trade such as silk were distributed across a large region, as bounded sets of laborers who shared an affiliation. These guilds played many roles, including training apprentices [18, 44], setting prices [45], and providing mechanisms for collective action [52, 49]. Especially relevant to the current challenge, guilds measured and certified their own members' quality [18]. While guilds eventually lost influence due to exerting overly tight controls on trade [45] and exogenous technical innovations in production, their intellectual successors persist today as professional organizations such as in engineering, acting and medicine [46, 33]. Malone first promoted a vision of online "e-lancer" guilds twenty years ago [40], but to date no concrete instantiations exist for a modern, online crowd work economy.

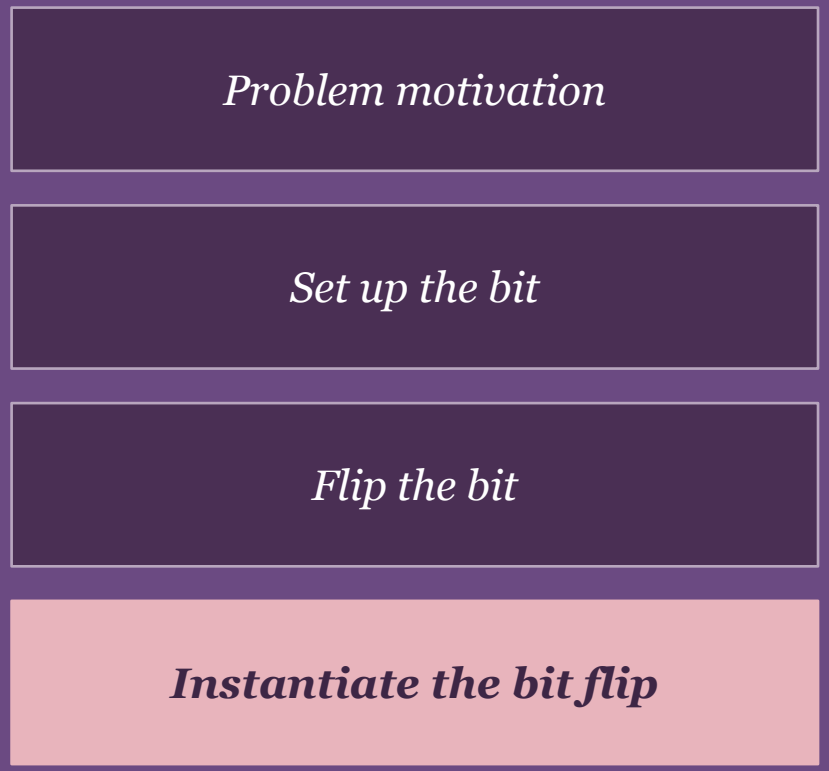
flip = re-centralize via guilds

The rest of the paragraph explains the high-level idea.

Instantiate the bit flip

At this point, the reader understands the idea that you're proposing, but it's still very high level. In this paragraph, map that idea onto a concrete instantiation.

Typically, this is where the system or algorithm that you're creating gets a name. Explain its architecture or design at a high level. Make clear how this architecture or design is an instance of the bit flip.

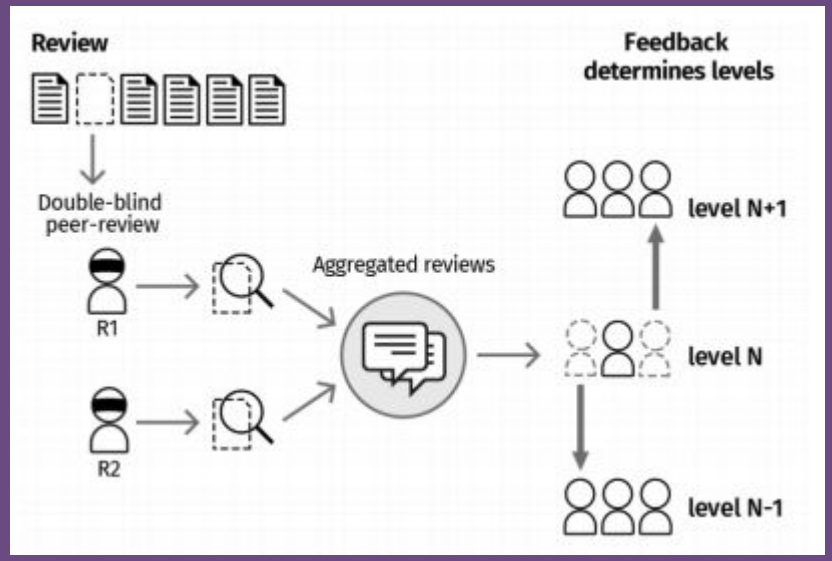


We present *crowd guilds*: crowd worker collectives that coordinate to certify their own members and perform internal feedback to train members (Figure 1). Our infrastructure for crowd guilds enables workers to engage in continuous double-blind peer assessment [30] of a random sample of members' task submissions on the crowdsourcing platform, rating the quality of the submission and providing critiques for further improvement. These peer assessments are used to derive guild levels (e.g., Level 1, Level 2) to serve as reputation (qualification) signals on the crowdsourcing platform. As workers gather positive assessments from more senior guild members, they rise in levels within the guild. Guilds translate these levels into higher wages by recommending pay rates for each level when tasks are posted to the platform. While crowd guilds focus here on worker reputation, our experiment implementation also explores how crowd guilds could address other challenges such as collective action (e.g., collectively rejecting tasks that pay too little), formal mentorship (e.g., repeated feedback and training), and social support (e.g., on the forums). Because

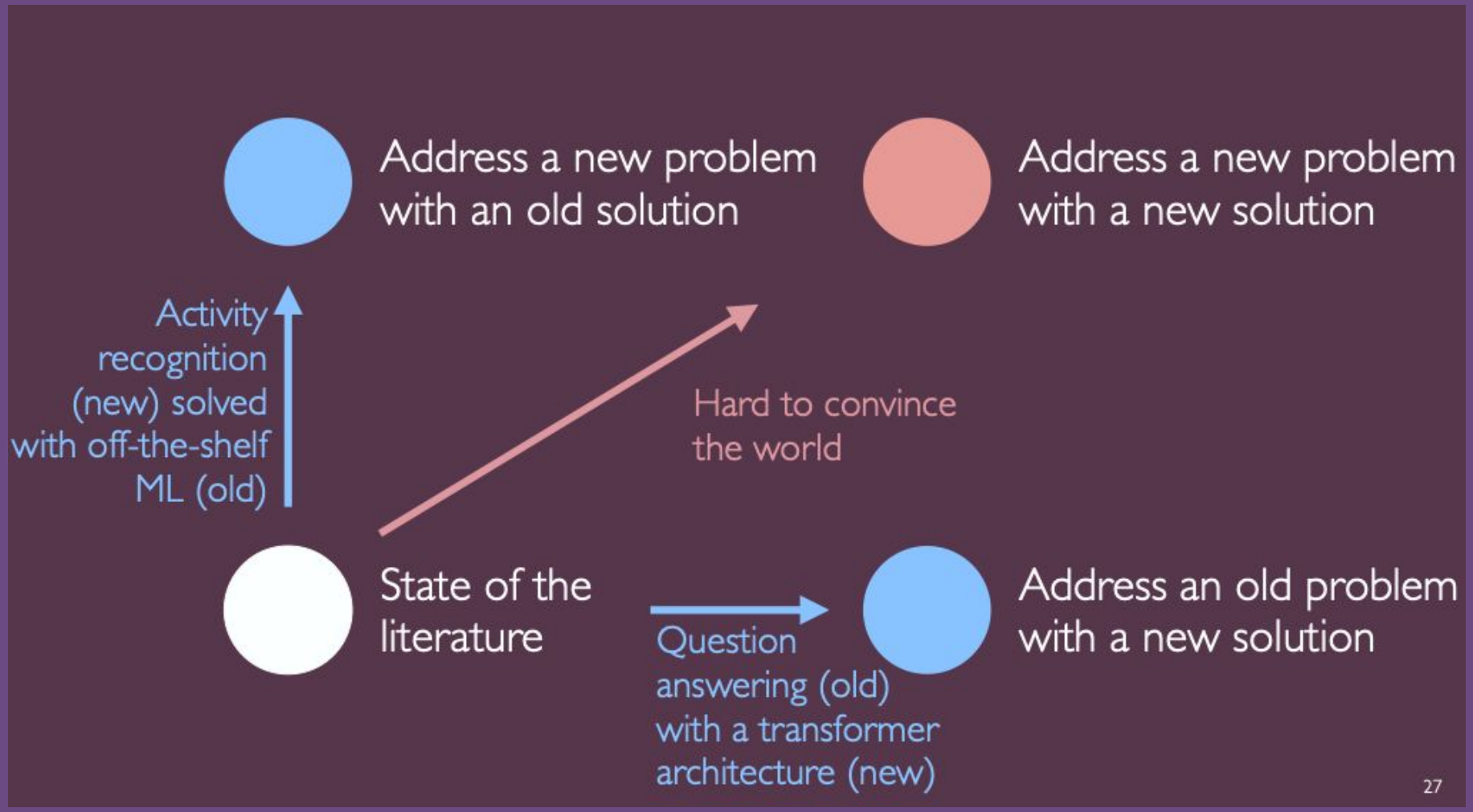
instantiation =

crowd guilds system

The rest of the paragraph details how crowd guilds work.



Types of bit flips



Why only make one move?

When making an argument, you want to introduce one major new idea, to minimize the new ideas your listener needs to absorb. A research paper typically only flips one bit.

Typically you are spending the introduction making the case for your new idea. If you are trying to make the case for both a new problem and a new solution, a reader might disagree with either.

This is not to say that you can't do new problem / new solution; just that it's a risky varsity maneuver.

Try again

Turn to your partner and explain your project. Cheat sheet on the right. [1 min each]

How clearly do you understand your partner's problem?

How clearly do you understand your partner's solution?

Problem motivation

Set up the bit

Flip the bit

Instantiate the bit flip

Evaluation and Implications

Evaluation: We'll talk about next.

Implications: If you're right and the bit flip is how everyone should be approaching this problem from now on, what implications are there for the field?

Will it change the contexts in which we use this technology? Will it broaden usage?