

CMSC 829C: Algorithms, AI, and Hardware for Biological Data

Lecture 2: Moving the Needle - A Bioinformatics Perspective

Fall 2026

September 3, 2026

Can Firtina | canfirtina@gmail.com | cs.umd.edu/~firtina

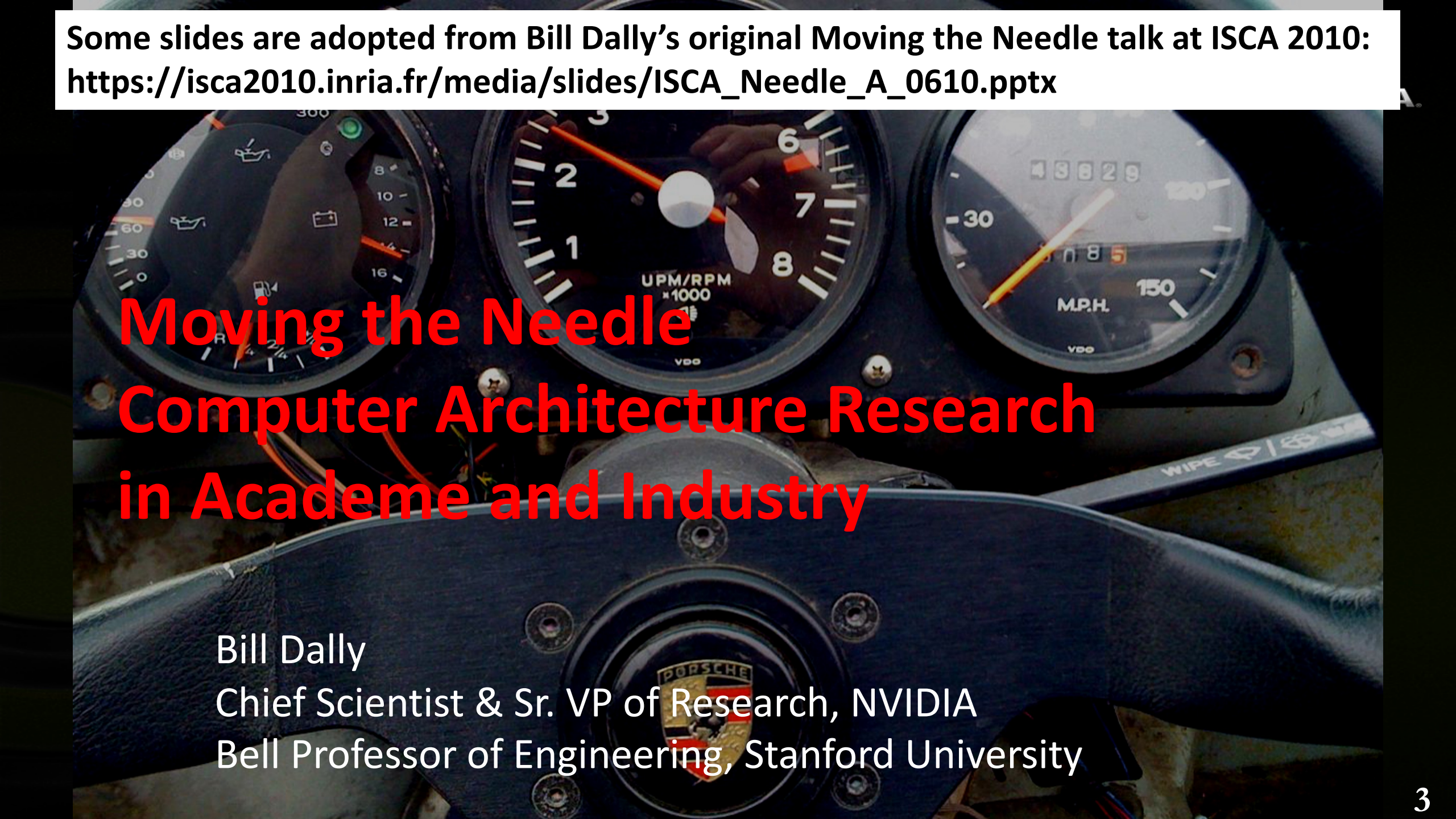
Assistant Professor of Computer Science at UMD



Recall: Some More Materials to Internalize

- Bill Dally,
["Moving the needle: Effective Computer Architecture Research in Academe and Industry"](#)
ISCA, 2010.
[\[Alternative Talk Video\]](#)
- Michael Stonebraker,
["Make it Happen"](#)
ACM Turing Award Lecture, 2014.
- Martin A. Schwartz,
["The importance of stupidity in scientific research"](#)
Journal of Cell Science, 2008.

Some slides are adopted from Bill Dally's original Moving the Needle talk at ISCA 2010:
https://isca2010.inria.fr/media/slides/ISCA_Needle_A_0610.pptx



Moving the Needle Computer Architecture Research in Academe and Industry

Bill Dally

Chief Scientist & Sr. VP of Research, NVIDIA

Bell Professor of Engineering, Stanford University

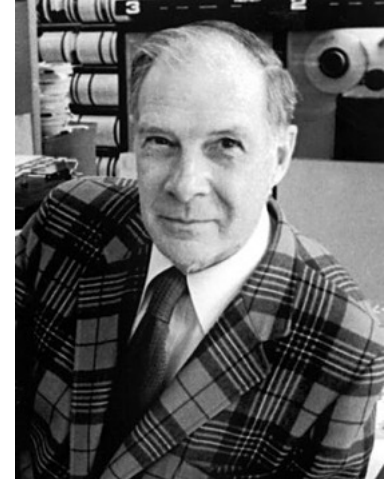
Research is a
hunt for insight



Recall: The Goal of Computing: Beyond Numbers

"The purpose of computing is insight, not numbers"

Richard Hamming



Recall: You and Your Research

- Richard Hamming,
["You and Your Research"](#)
Bell Communications Research Colloquium Seminar, 1986.
[\[Lecture Video\]](#) [\[Podcast\]](#)

Richard Hamming

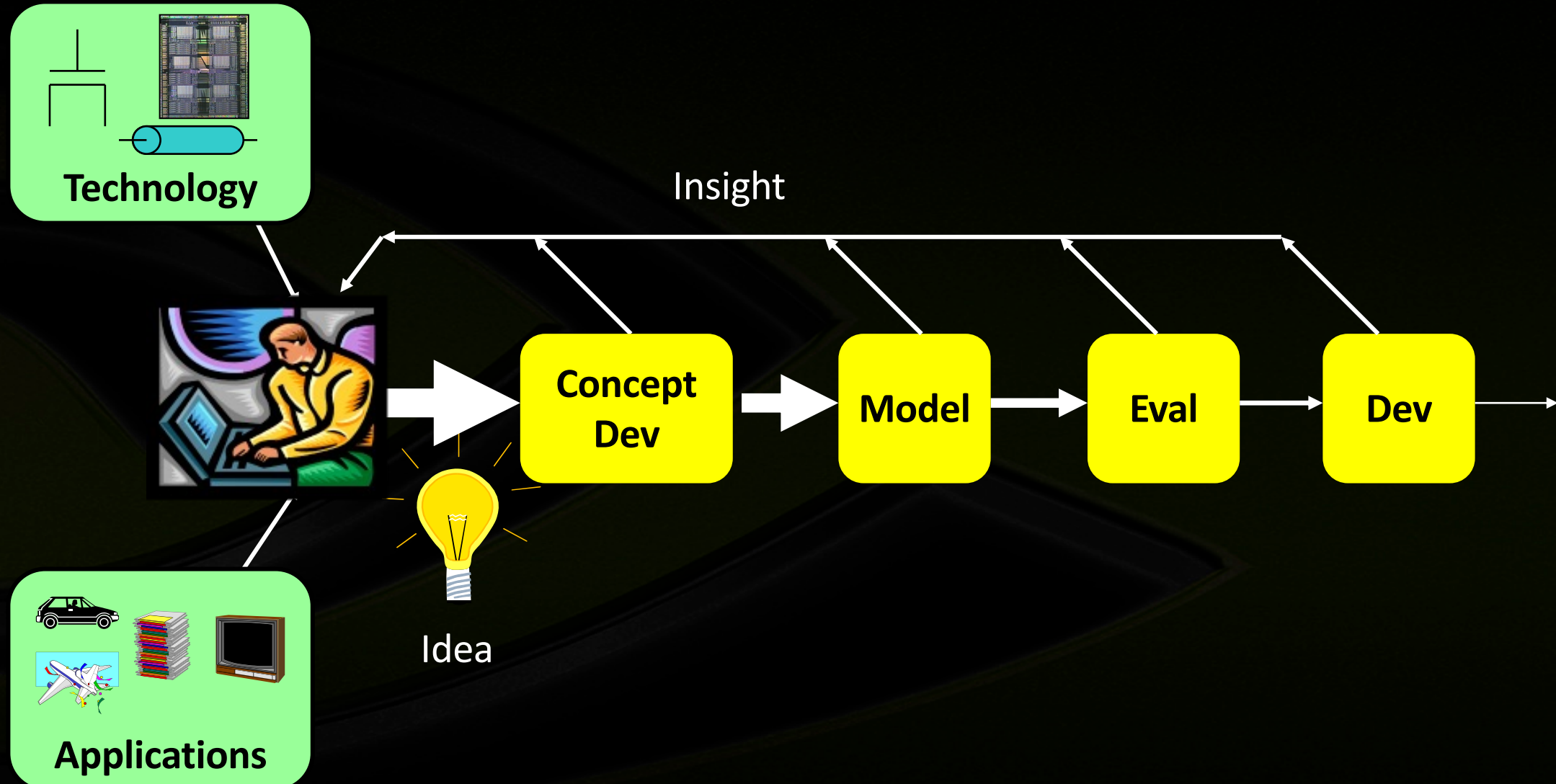
``You and Your Research''

Transcription of the
Bell Communications Research Colloquium Seminar
7 March 1986

Goal – Positive Impact ~~on a Product~~

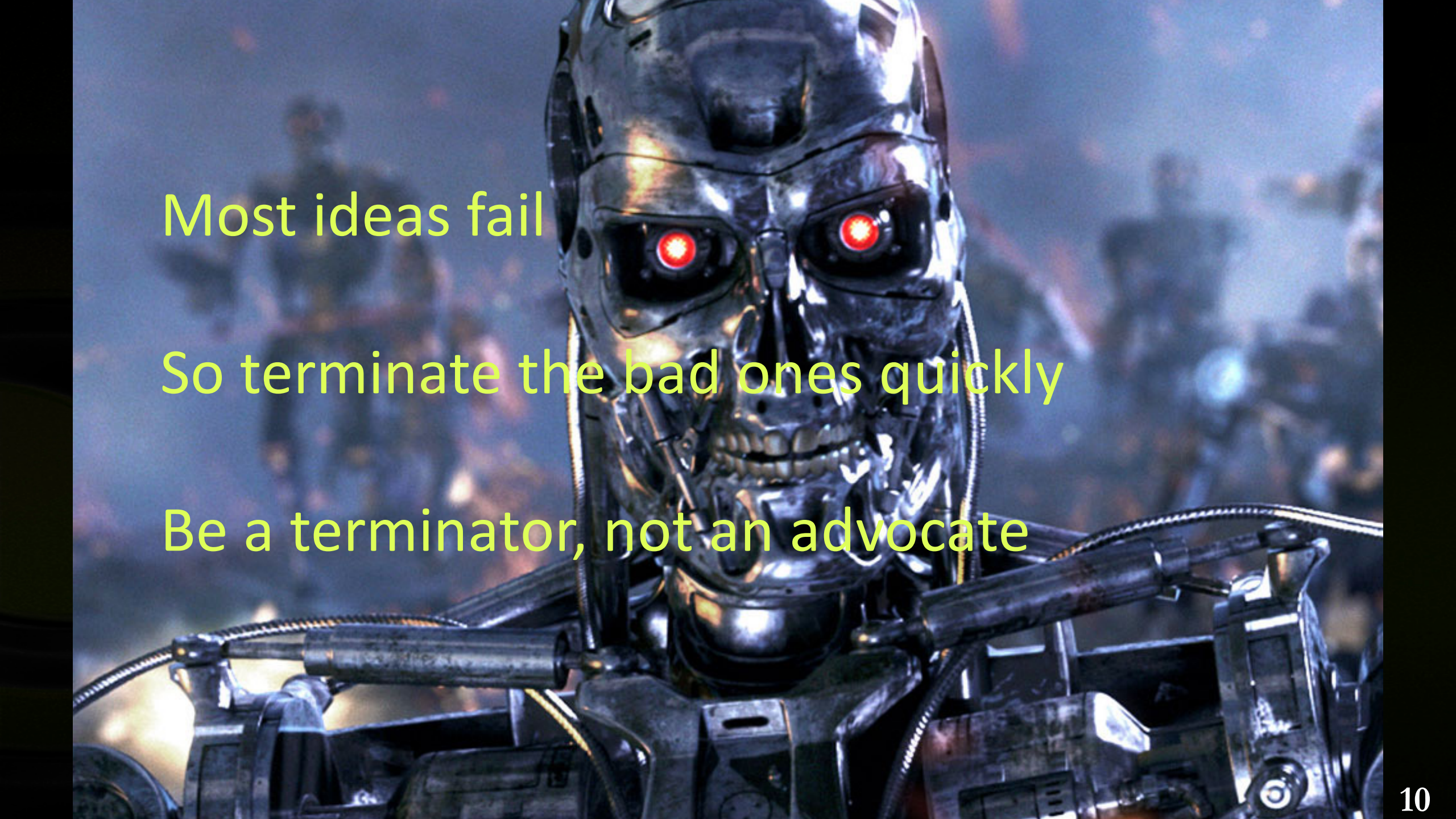


The Research Funnel



Most ideas fail

So terminate the bad ones quickly



Most ideas fail

So terminate the bad ones quickly

Be a terminator, not an advocate

Heilmeier's Catechism

- What are you trying to do? Articulate your objectives using absolutely no jargon.
- How is it done today, and what are the limits of current practice?
- What's new in your approach and why do you think it will be successful?
- Who cares? If you're successful, what difference will it make?
- What are the risks and the payoffs?
- How much will it cost?
- How long will it take?
- What are the midterm and final "exams" to check for success?
- <https://techtransferfordefense.stanford.edu/creating-impactful-one-pagers-using-heilmeier-catechism>

How to Present Papers/Ideas/Projects

- Background
- Problem
 - What are the limitations of the state-of-the-art approach(es)
- Goal
 - Articulate what the paper is trying to achieve with absolutely no jargons
- Key Idea/Mechanism
 - What are the idea(s) and steps to achieve this goal
- Key (Expected) Results
 - What are the key observations?
 - What insights do we generate from them?
- Strengths and Weaknesses
- Future Directions and Ideas

New Frontiers: Raw Signal Analysis [ISMB '23]

- **RawHash: enabling fast and accurate real-time analysis of raw nanopore signals for large genomes**

Can Firtina, Nika Mansouri Ghiasi, Joel Lindegger, Gagandeep Singh, Meryem Banu Cavlak, Haiyu Mao, Onur Mutlu

[Proceedings of the 31st Annual Conference on Intelligent Systems for Molecular Biology and the 22nd European Conference on Computational Biology \(ISMB/ECCB 2023\)](#), Lyon, France, July 2023.

[\[PDF\]](#) [\[Code\]](#) [\[DOI\]](#)

Conference talk [\[Video\]](#) [\[Slides \(pptx\)\]](#) [\[Slides \(pdf\)\]](#)

Preprint (bioRxiv) [\[Link\]](#) [\[PDF\]](#)

Social media thread [\[Twitter \(X\)\]](#)

Bioinformatics, 2023, **39**, i297–i307
<https://doi.org/10.1093/bioinformatics/btad272>
ISMB/ECCB 2023

OXFORD

RawHash: enabling fast and accurate real-time analysis of raw nanopore signals for large genomes

Can Firtina ^{1,*}, Nika Mansouri Ghiasi ¹, Joel Lindegger ¹, Gagandeep Singh ¹,
Meryem Banu Cavlak ¹, Haiyu Mao ¹, Onur Mutlu ^{1,*}

¹Department of Information Technology and Electrical Engineering, ETH Zurich, 8092 Zurich, Switzerland

*Corresponding author. Department of Information Technology and Electrical Engineering, ETH Zurich, Gloriastrasse 35, 8092 Zurich, Switzerland.
E-mail: firtinac@ethz.ch (C.F.), omutlu@ethz.ch (O.M.)

How to Present Papers/Ideas/Projects

- **Background**
- Problem
 - What are the limitations of the state-of-the-art approach(es)
- Goal
 - Articulate what the paper is trying to achieve with absolutely no jargons
- Key Idea/Mechanism
 - What are the idea(s) and steps to achieve this goal
- Key (Expected) Results
 - What are the key observations?
 - What insights do we generate from them?
- Strengths and Weaknesses
- Future Directions and Ideas

Executive Summary

Problem: Real-time analysis of nanopore raw signals is **inaccurate** and **inefficient** for **large genomes**

Goal: Enable **fast** and **accurate** real-time analysis of raw signals for **large genomes**

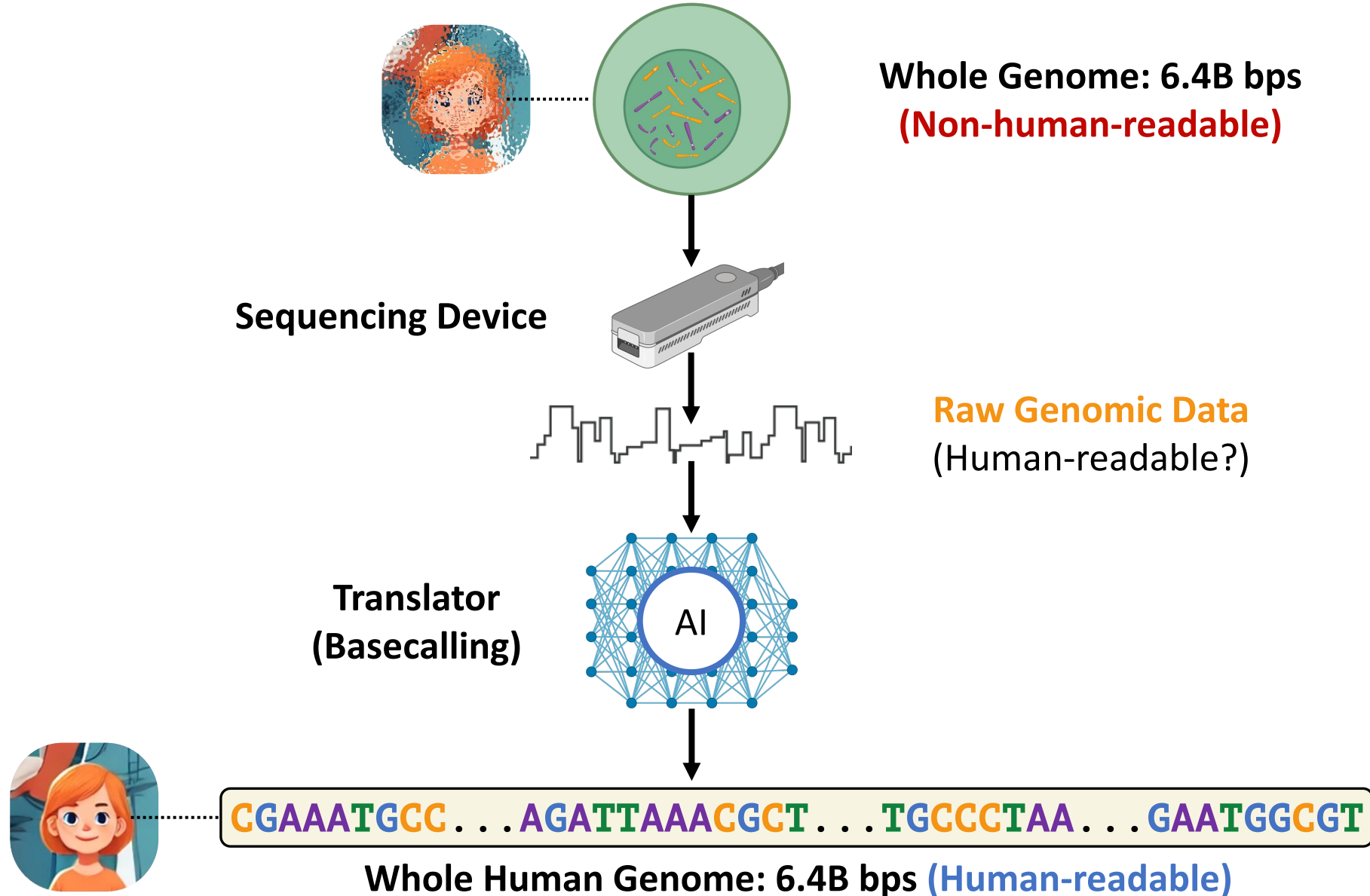
Key Contributions:

- 1) The **first hash-based mechanism** that can quickly and accurately analyze raw nanopore signals for **large genomes**
- 2) The novel **Sequence Until** technique can accurately and **dynamically stop the entire sequencing of all reads at once** if further sequencing is not necessary

Key Results: Across 3 use cases and 5 genomes of varying sizes, RawHash provides

- **25.8× and 3.4× better average throughput** compared to two state-of-the-art works
- **1.14× – 2.13× more accurate mapping results** for **large genomes**
- **Sequence Until reduces the sequencing time and cost by 15×**

Recall: Generating the Entire Genomic Sequence

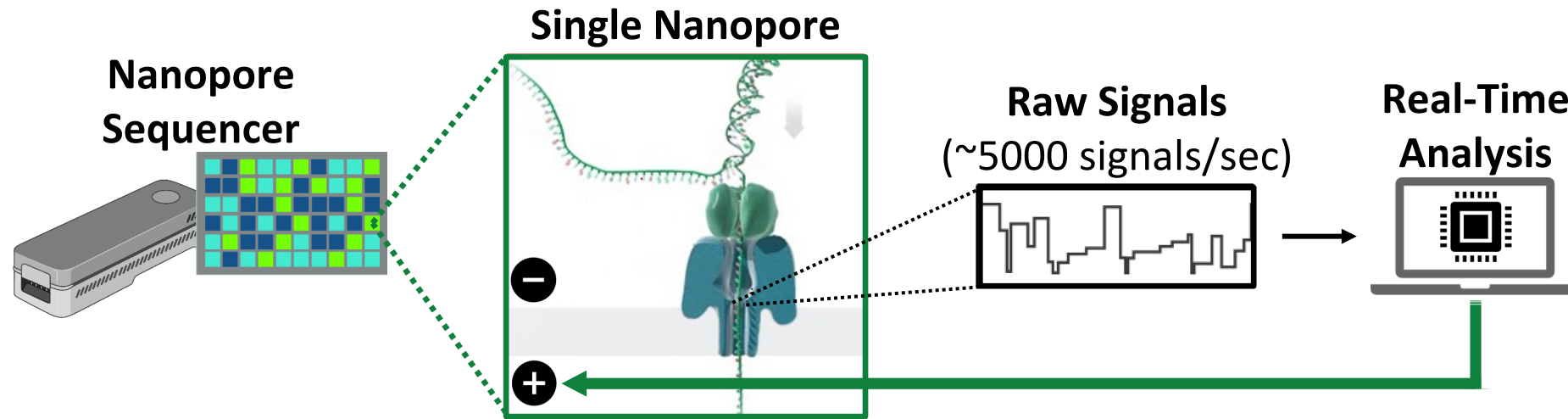


Nanopore Sequencing

- Nanopore Sequencing: a widely used sequencing technology
 - Can sequence large fragments of nucleic acid molecules (up to >2Mbp)
 - Offers high throughput
 - Cost-effective
 - Enables real-time genome analysis



Generating Electrical Signals from DNA



Raw Signals: Ionic current measurements generated at a certain **throughput**

Real-Time Analysis: Analyzing raw signals **instantly** as they are generated

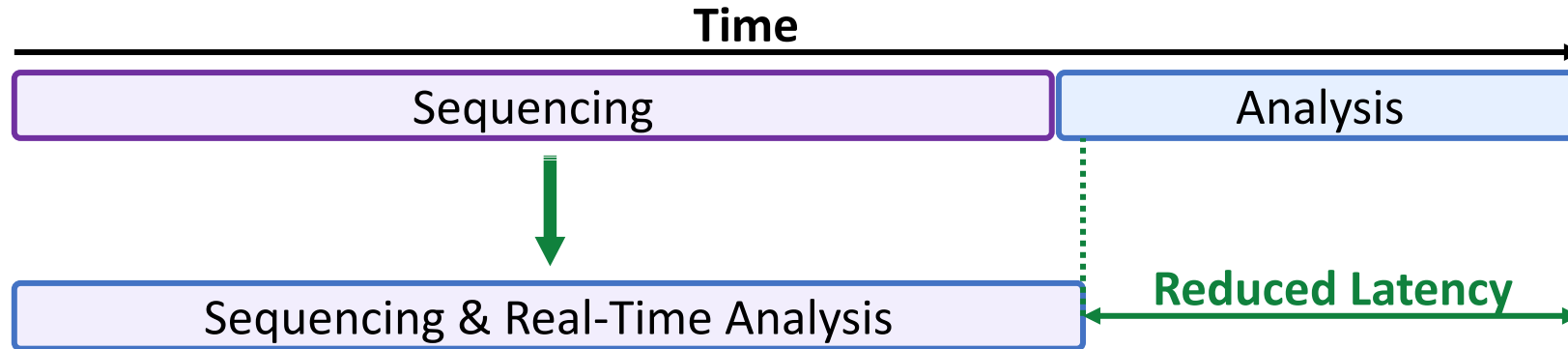
Real-Time Decisions: Stopping sequencing **early** based on real-time analysis

<https://nanoporetech.com/platform/technology/basecalling>

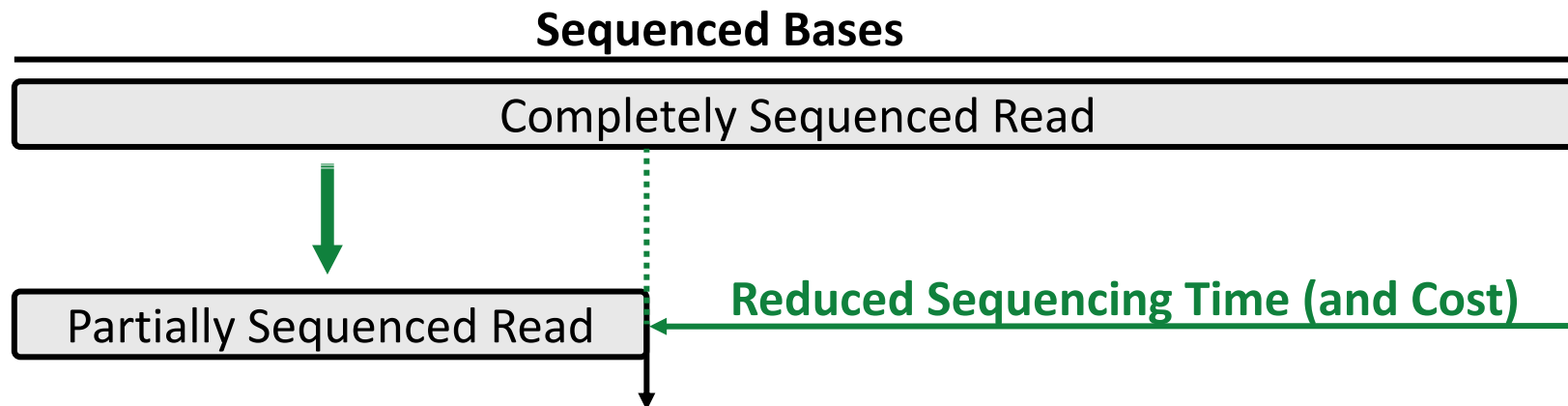
<https://blogs.ed.ac.uk/institute-genetics-cancer/2020/04/08/nanopore-sequencing-controversial-world-records-and-whale-watching/>

Benefits of Real-Time Analysis

- ✓ **Reducing latency** by overlapping analysis with sequencing

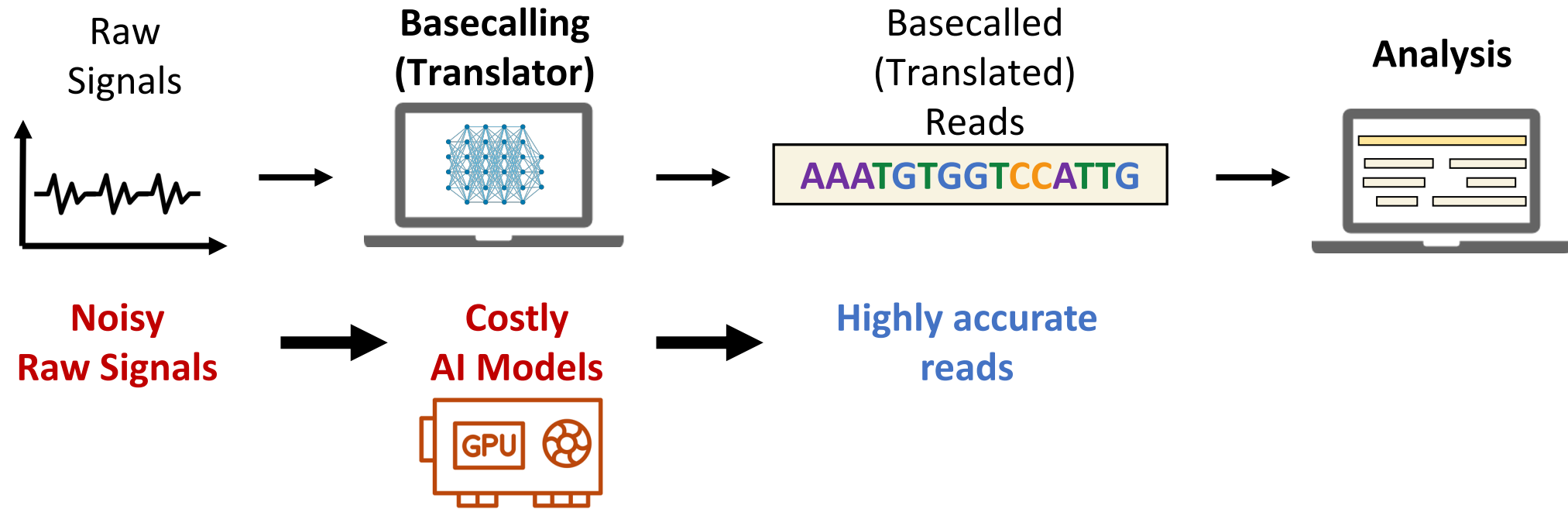


- ✓ **Reducing sequencing time and cost** by stopping sequencing early

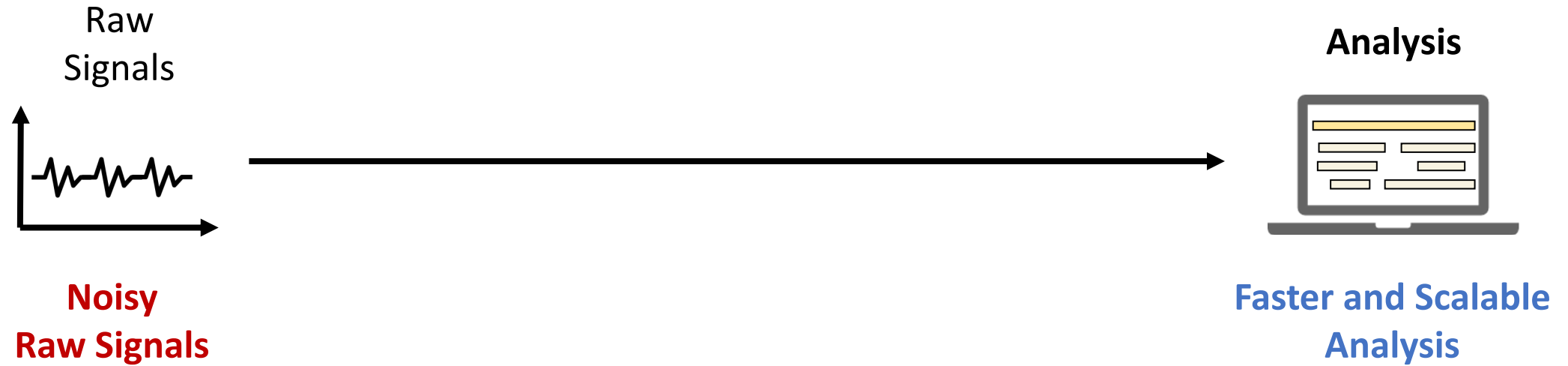


Sequencing is stopped early with a real-time decision

A Common Genome Analysis Pipeline



Alternative: Directly Analyze Raw Signals



Noise in Raw Electrical Signals



Observation: Identical molecules generate similar raw signals

Sequencing CTGCGTAGCA

Raw Signals: [-0.06, -0.22, -0.75, 0.96, -0.49]

Prior Work: Perform vector similarity estimation based on nearest neighbor search

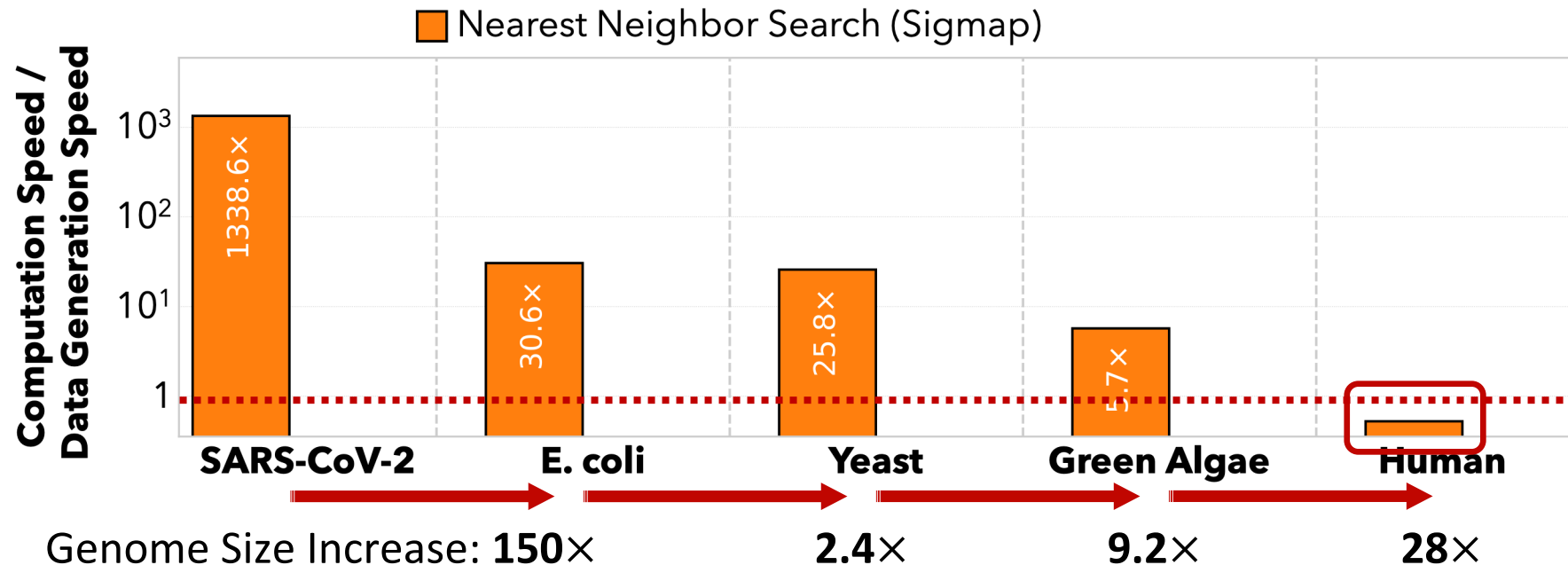
Raw Signals: [-0.05, -0.21, -0.76, 1.01, -0.51]

Sequencing CTGCGTAGCA

How to Present Papers/Ideas/Projects

- Background
- **Problem**
 - **What are the limitations of the state-of-the-art approach(es)**
- Goal
 - Articulate what the paper is trying to achieve with absolutely no jargons
- Key Idea/Mechanism
 - What are the idea(s) and steps to achieve this goal
- Key (Expected) Results
 - What are the key observations?
 - What insights do we generate from them?
- Strengths and Weaknesses
- Future Directions and Ideas

Scalability Issues Limit Real-Time Analysis



Problem: No Real-Time Analysis for Larger Genomes:
Searching mechanism does not scale well

How to Present Papers/Ideas/Projects

- Background
- Problem
 - What are the limitations of the state-of-the-art approach(es)
- **Goal**
 - **Articulate what the paper is trying to achieve with absolutely no jargons**
- Key Idea/Mechanism
 - What are the idea(s) and steps to achieve this goal
- Key (Expected) Results
 - What are the key observations?
 - What insights do we generate from them?
- Strengths and Weaknesses
- Future Directions and Ideas

RawHash's Goal

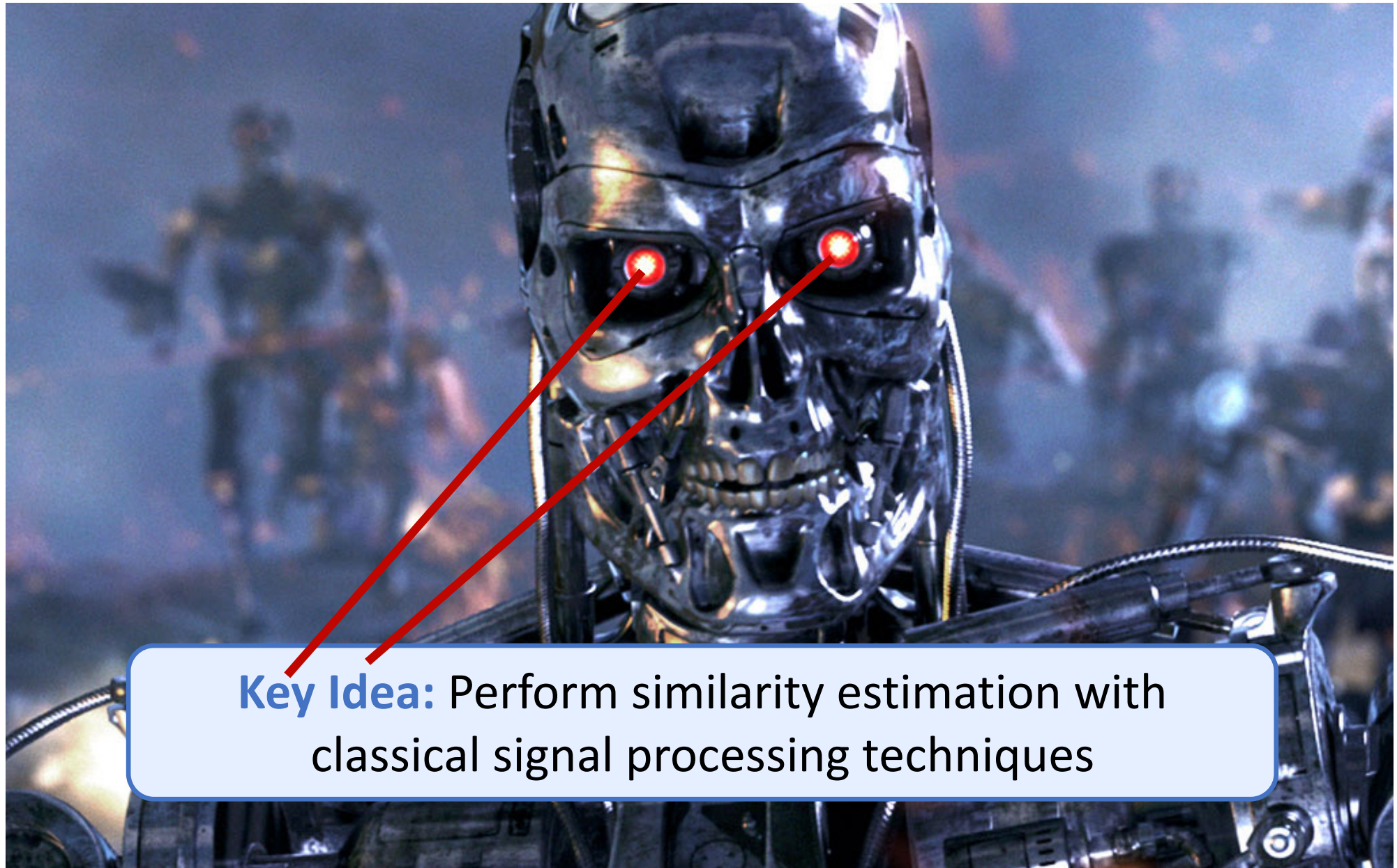
Enable **Fast, Accurate, and Scalable**
Real-Time Analysis for Larger Genomes

How to Present Papers/Ideas/Projects

- Background
- Problem
 - What are the limitations of the state-of-the-art approach(es)
- Goal
 - Articulate what the paper is trying to achieve with absolutely no jargons
- **Key Idea/Mechanism**
 - **What are the idea(s) and steps to achieve this goal**
- Key (Expected) Results
 - What are the key observations?
 - What insights do we generate from them?
- Strengths and Weaknesses
- Future Directions and Ideas

A Key Idea

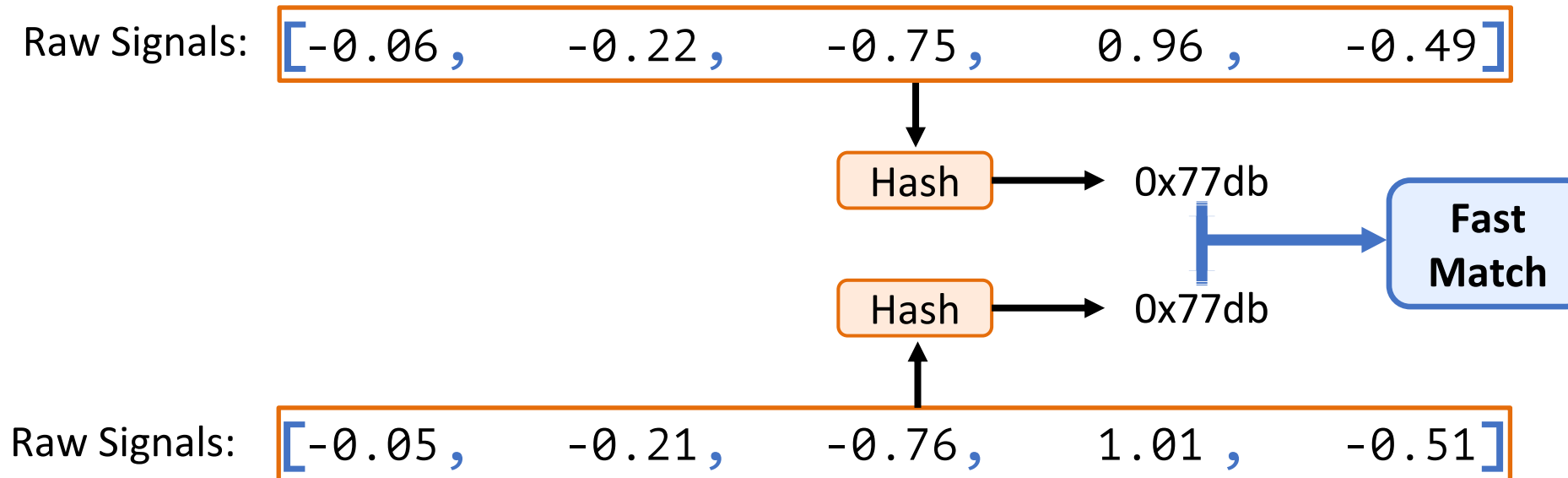
Key Idea: Perform similarity estimation with signal processing techniques



Key Idea: Perform similarity estimation with classical signal processing techniques

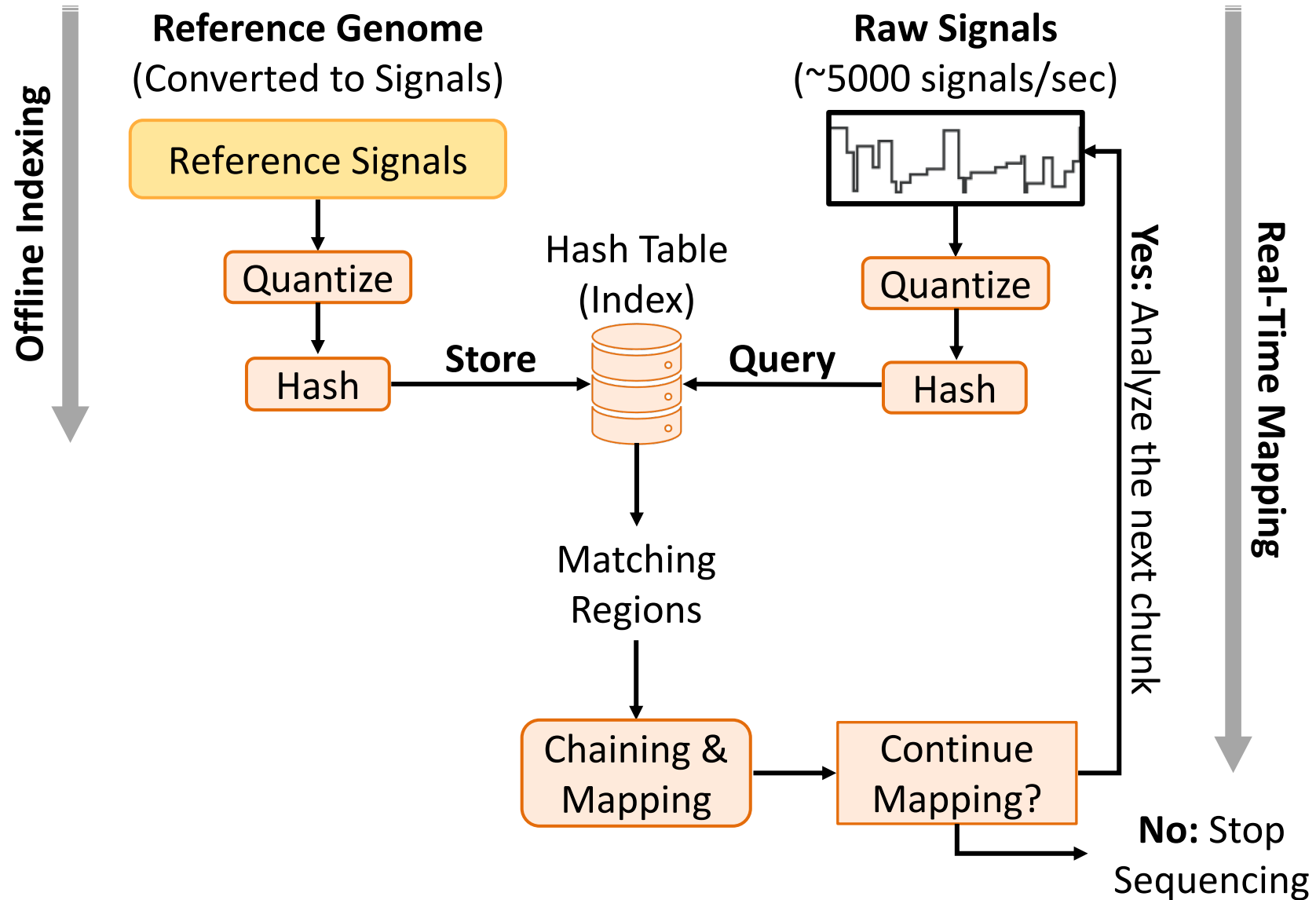
RawHash's Key Idea

Key Idea: Perform similarity estimation based on **quick exact matches**



Challenge: How to generate **the same hash value** for **similar enough signals with various noise types**?

RawHash Overview



Step 1: Quantization

Reference Genome
(Converted to Signals)

Reference Signals

Quantize

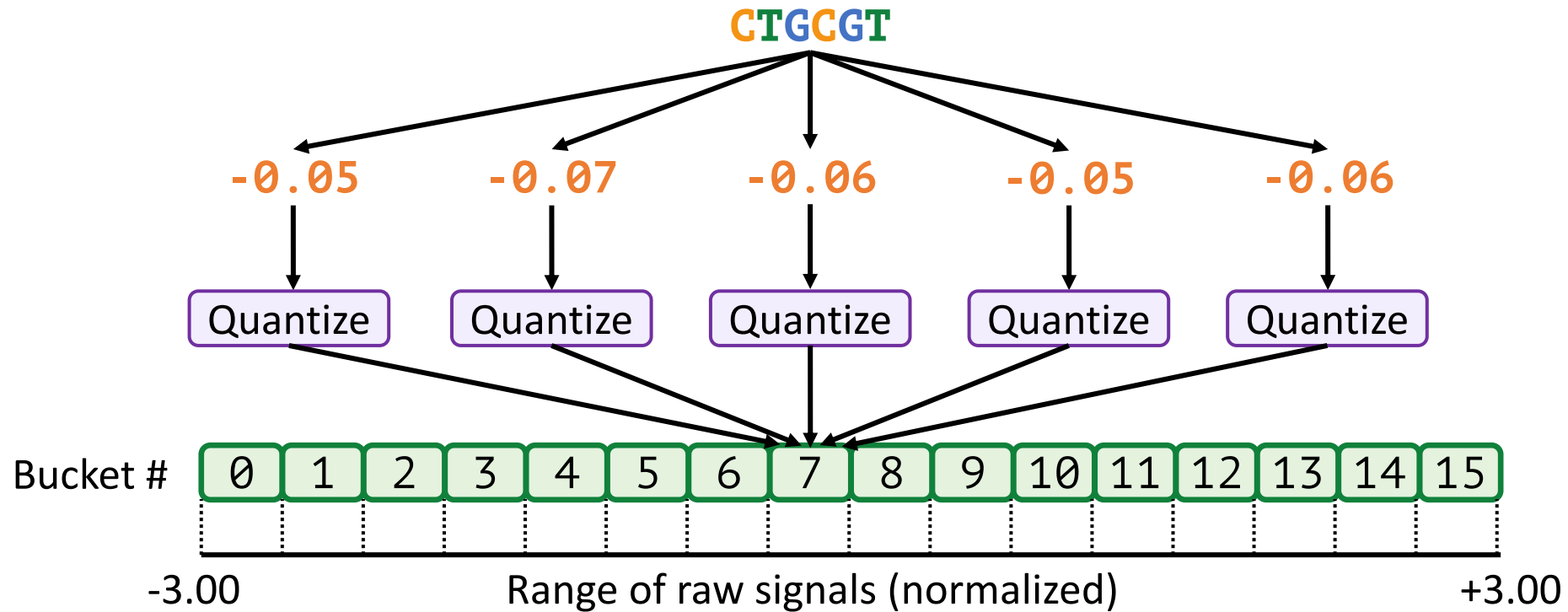
Raw Signals
(~5000 signals/sec)



Quantize

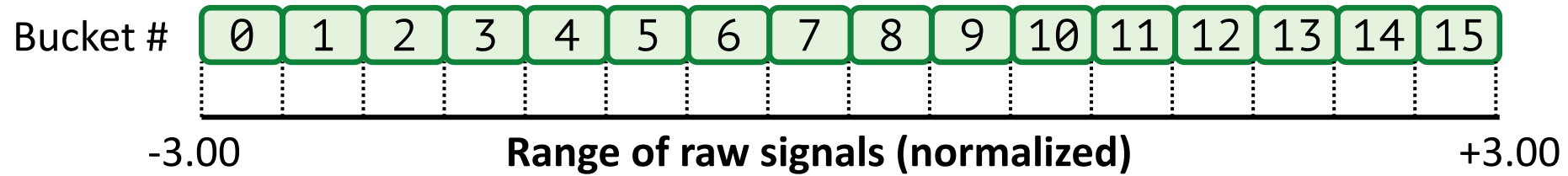


RawHash's Key Idea – Quantization

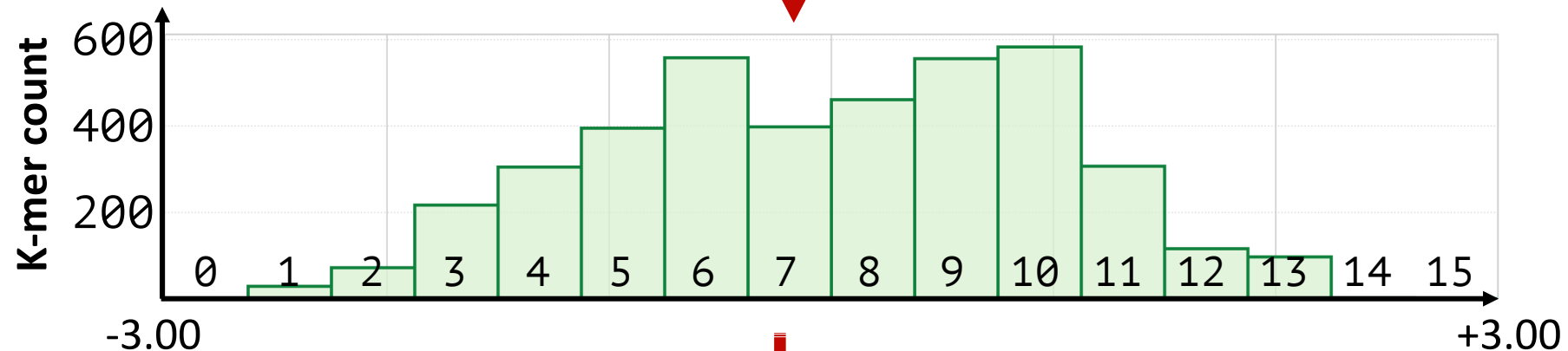


Enables matching raw signals by eliminating slight differences

How to Better Quantize Signals



✗ Equal-width buckets leads to **unbalanced bucket loading**

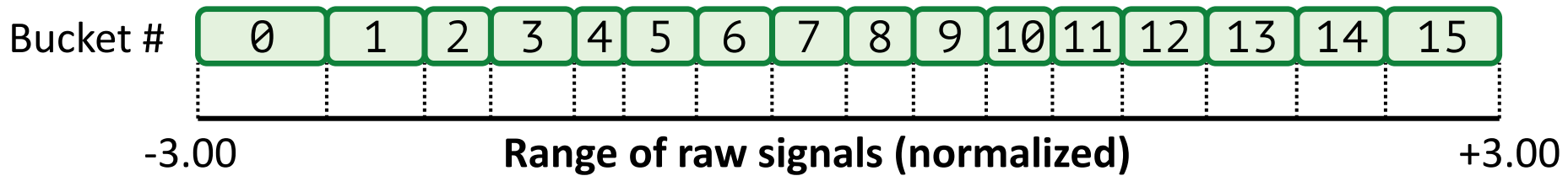


✗ Reduced uniqueness and poor resource utilization

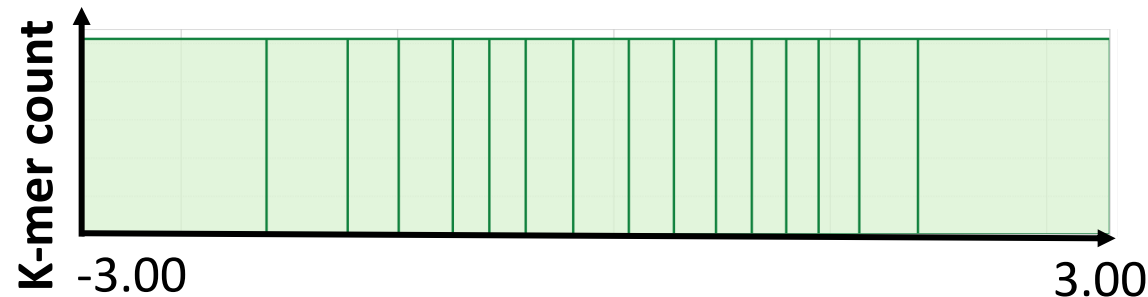


Adaptive Quantization

- **Key Idea:** Quantizing raw signals with **non-equal bucket widths** by leveraging raw signal distribution

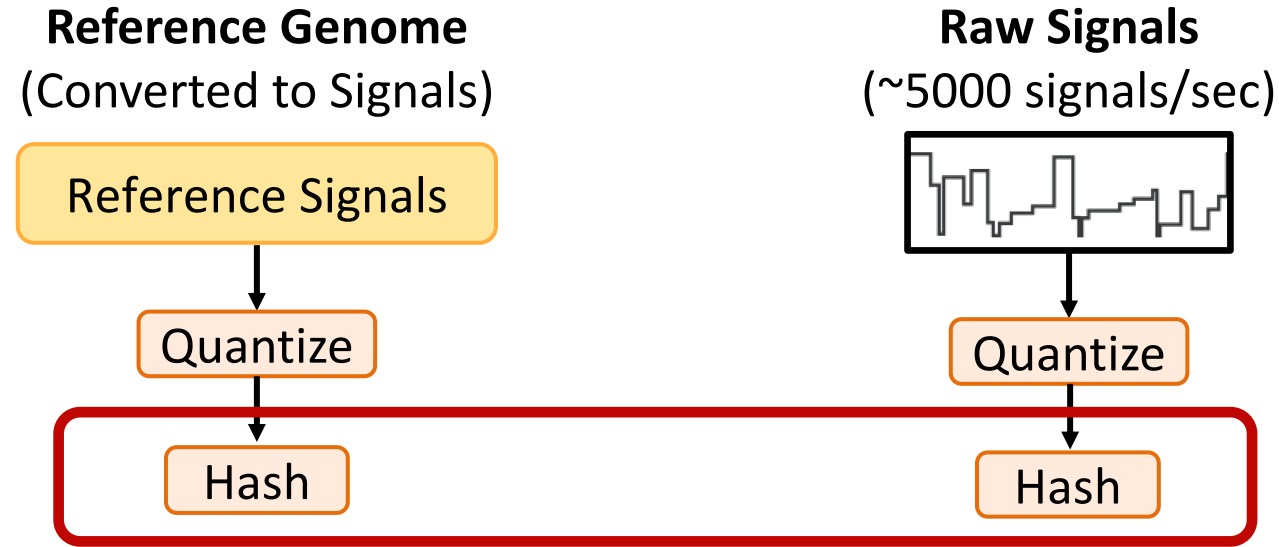


$$q(s) = \begin{cases} \left\lfloor n \times \left(f_r \times \frac{(s - f_{min})}{f_{max} - f_{min}} \right) \right\rfloor & \text{if } f_{min} \leq s \leq f_{max} \\ \left\lfloor n \times (f_r + c_r \times s) \right\rfloor & \text{if } s < f_{min} \\ \left\lfloor n \times (f_r + c_r + c_r \times s) \right\rfloor & \text{if } s > f_{max} \end{cases}$$

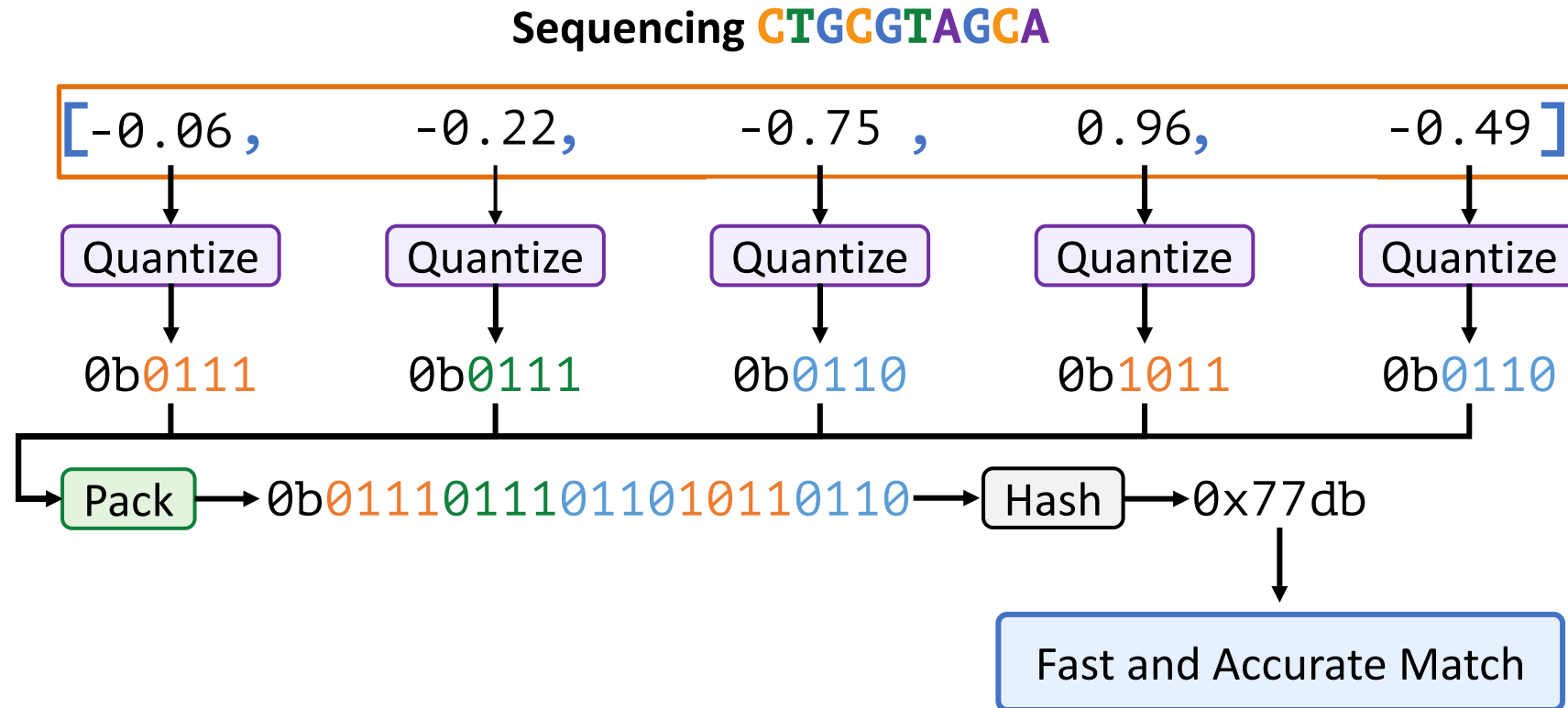


Better bucket utilization

Step 2: Hashing Quantized Signals

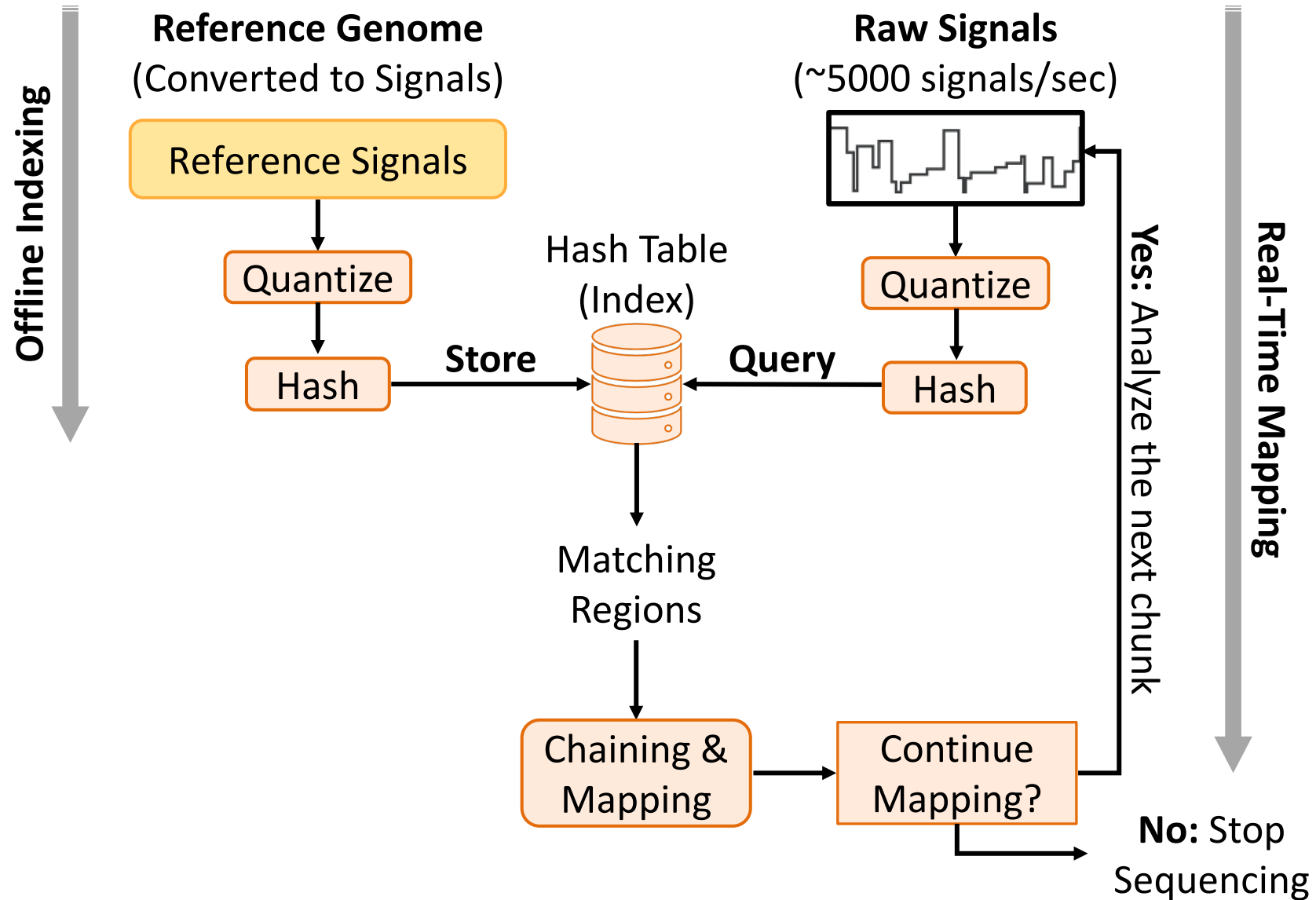


RawHash Key Idea – Hash-based Matching



Key Contribution: **The first** hash-based search (and indexing) mechanism for raw signals

RawHash Overview



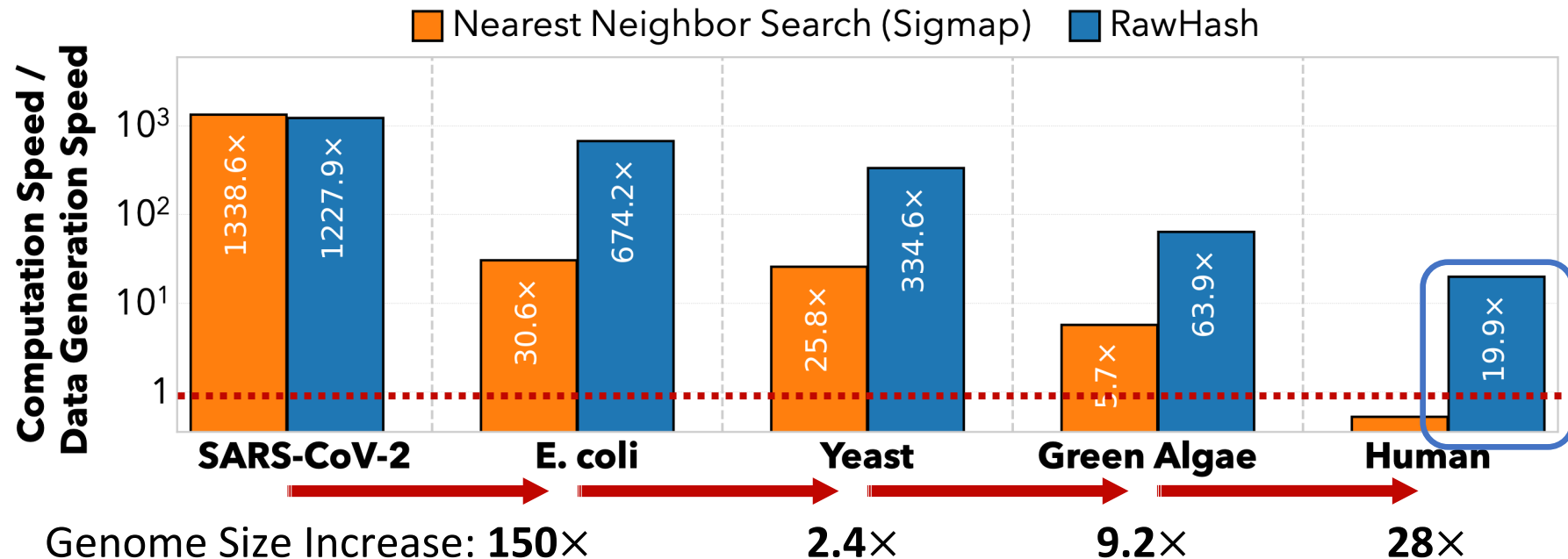
How to Present Papers/Ideas/Projects

- Background
- Problem
 - What are the limitations of the state-of-the-art approach(es)
- Goal
 - Articulate what the paper is trying to achieve with absolutely no jargons
- Key Idea/Mechanism
 - What are the idea(s) and steps to achieve this goal
- **Key (Expected) Results**
 - **What are the key observations?**
 - **What insights do we generate from them?**
- Strengths and Weaknesses
- Future Directions and Ideas

Evaluation Methodology

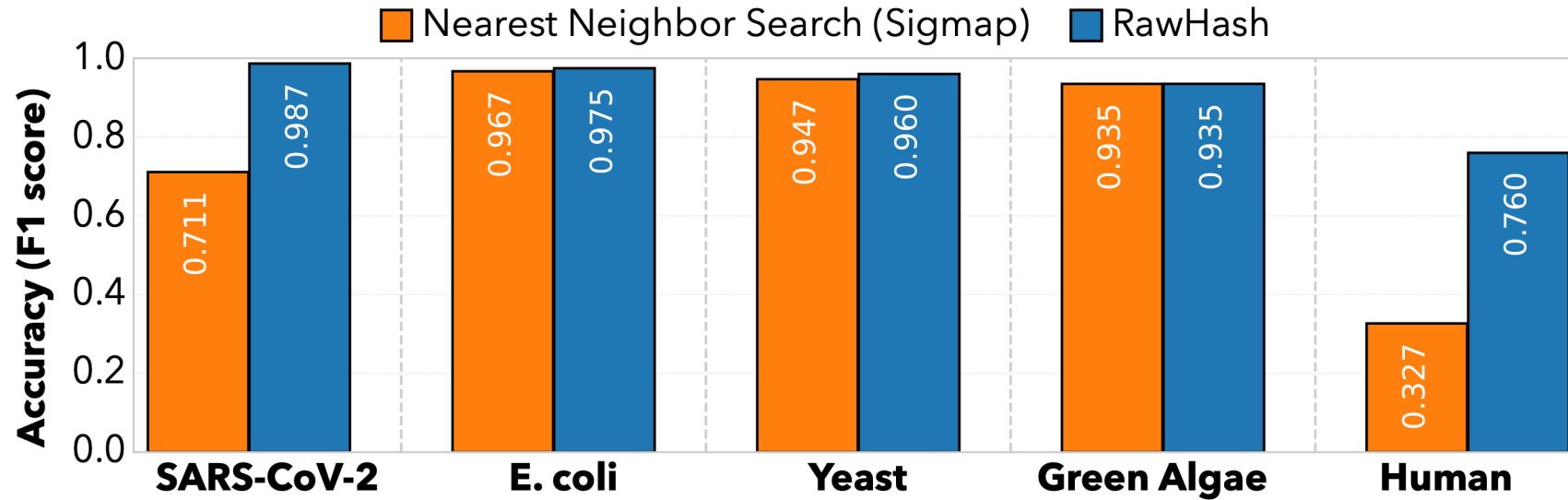
- Compared to **Sigmap** [Zhang+, ISMB/ECCB'21]
- **Evaluation metrics:**
 1. Throughput (computation speed / data generation speed)
 2. Read mapping accuracy (F1)
 - Ground truth: Basecalled read mapping
- You can also mention: Datasets, simulators, baseline implementation, hardware and so on
 - Evaluation methodology is where you describe the experimental setup used for generating the results you are about to show

Real-Time Analysis Benefits from Fast Search



Real-time Analysis for Human Genomes:
Fast search instead of costly neighbor search

Fast Search Leads to Accurate Search



Best accuracy for all genomes:

Hash values span longer sequences than vectors ->
More unique and informative matches

(Almost) Identical accuracy to basecalled analysis:

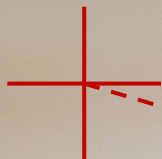
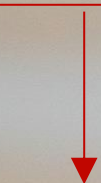
Is the small difference because RawHash is even better?



How to Present Papers/Ideas/Projects

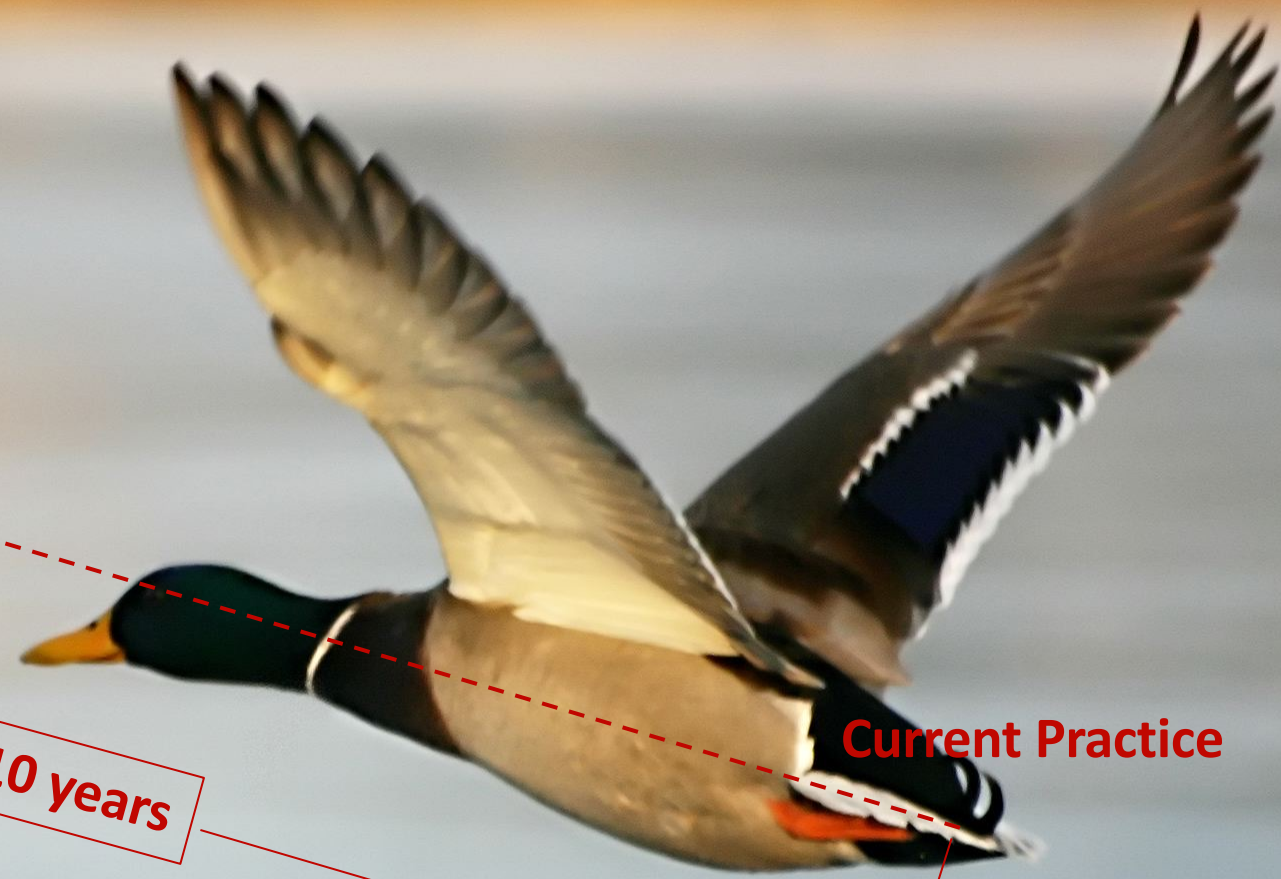
- Background
- Problem
 - What are the limitations of the state-of-the-art approach(es)
- Goal
 - Articulate what the paper is trying to achieve with absolutely no jargons
- Key Idea/Mechanism
 - What are the idea(s) and steps to achieve this goal
- Key (Expected) Results
 - What are the key observations?
 - What insights do we generate from them?
- Strengths and Weaknesses
- **Future Directions and Ideas**

Aim Here



5-10 years

Current Practice



In-Storage Raw Signal Analysis [ICS '25]

- **MARS: Processing-In-Memory Acceleration of Raw Signal Genome Analysis Inside the Storage Subsystem**

Melina Soysal, Konstantina Koliogeorgi, [Can Firtina](#), Nika Mansouri Ghiasi, Rakesh Nadig, Haiyu Mao, Geraldo Francisco Oliveira Junior, Yu Liang, Klea Zambaku, Mohammad Sadrosadati, Onur Mutlu

[Proceedings of the 39th International Conference on Supercomputing \(ICS 2025\)](#), Salt Lake City, Utah, USA, June 2025.

[\[PDF\]](#) [\[DOI\]](#)

Conference talk [\[Video\]](#) [\[Slides \(pptx\)\]](#) [\[Slides \(pdf\)\]](#)

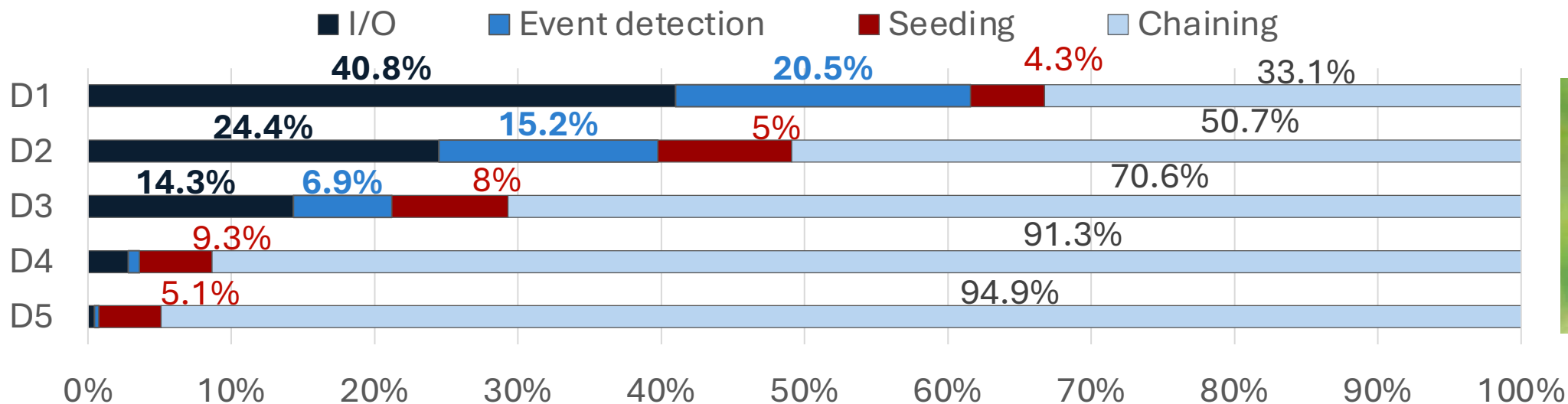
MARS: Processing-In-Memory Acceleration of Raw Signal Genome Analysis Inside the Storage Subsystem

Melina Soysal[†] Konstantina Koliogeorgi[†] Can Firtina[†] Nika Mansouri Ghiasi[†]
Rakesh Nadig[†] Haiyu Mayo^{*} Geraldo F. Oliveira[†]
Yu Liang[†] Klea Zambaku[†] Mohammad Sadrosadati[†] Onur Mutlu[†]

[†] *ETH Zürich*

^{*} *King's College London*

Motivation

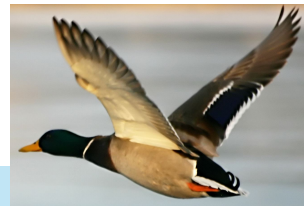
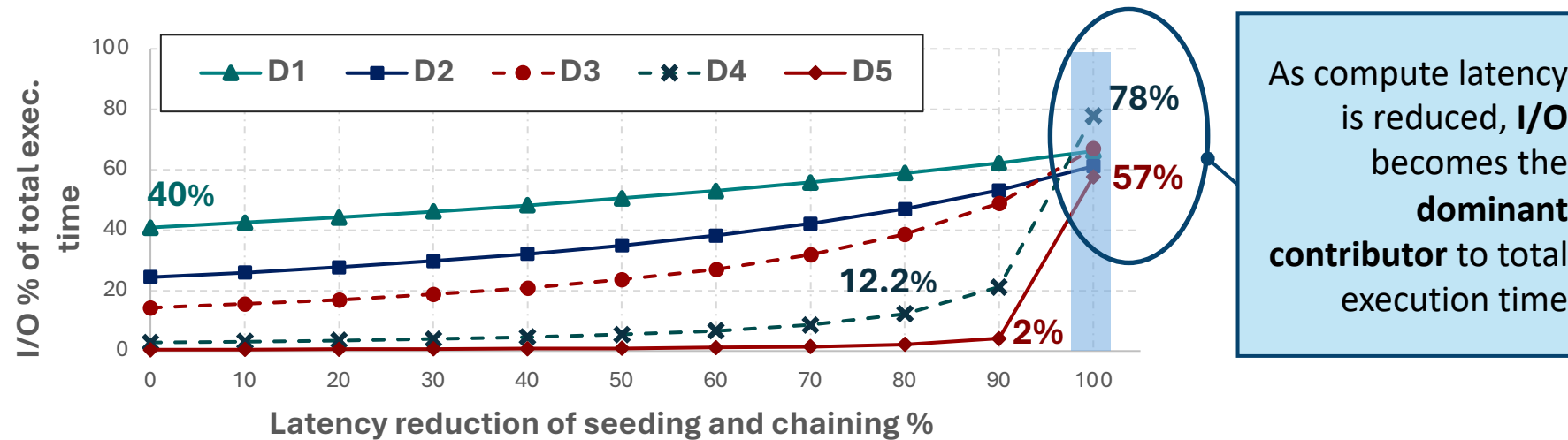


Chaining & **seeding** consistently dominate runtime

Event detection & **I/O overhead** are non-negligible for some datasets

Impact of I/O data movement

We model **I/O data movement overhead** when accelerating commonly targeted steps in the pipeline: **seeding** and **chaining**



I/O data movement from SSD becomes dominant overhead as compute steps are accelerated in RSGA

In-Storage Raw Signal Analysis [ICS '25]

- **MARS: Processing-In-Memory Acceleration of Raw Signal Genome Analysis Inside the Storage Subsystem**

Melina Soysal, Konstantina Koliogeorgi, [Can Firtina](#), Nika Mansouri Ghiasi, Rakesh Nadig, Haiyu Mao, Geraldo Francisco Oliveira Junior, Yu Liang, Klea Zambaku, Mohammad Sadrosadati, Onur Mutlu

[Proceedings of the 39th International Conference on Supercomputing \(ICS 2025\)](#), Salt Lake City, Utah, USA, June 2025.

[\[PDF\]](#) [\[DOI\]](#)

Conference talk [\[Video\]](#) [\[Slides \(pptx\)\]](#) [\[Slides \(pdf\)\]](#)

MARS: Processing-In-Memory Acceleration of Raw Signal Genome Analysis Inside the Storage Subsystem

Melina Soysal[†] Konstantina Koliogeorgi[†] Can Firtina[†] Nika Mansouri Ghiasi[†]
Rakesh Nadig[†] Haiyu Mayo^{*} Geraldo F. Oliveira[†]
Yu Liang[†] Klea Zambaku[†] Mohammad Sadrosadati[†] Onur Mutlu[†]

[†] *ETH Zürich*

^{*} *King's College London*

The contribution is *insight*

Not novelty

Not numbers

Research is a
hunt for insight

Need to get off the beaten
path to find new insights



Lost keys here



Looking here for lost keys



Conclusion

- Research is a multi-step process (art + science)
 - Focus on important problems
 - Figure out a way to test your ideas quickly
 - Build resilience
- As researchers, we seek insight to solve important problems
- Communicate the insights you generate extremely clearly
 - Otherwise, those insights will need to be rediscovered
- Keep your mind open

Questions?

Next Lecture

- **Lecture 3: Introduction to Genomics and Sequencing**
 - How sequencing technologies work and initial steps of its analysis
- **Before next class**
 - Finalize your paper selection (due Sep 8, 11:59pm)

CMSC 829C: Algorithms, AI, and Hardware for Biological Data

Lecture 2: Moving the Needle - A Bioinformatics Perspective

Fall 2026

September 3, 2026

Can Firtina | canfirtina@gmail.com | cs.umd.edu/~firtina

Assistant Professor of Computer Science at UMD

