

## Luis Gravano

- Columbia 1997 ~
- Stanford 1994~1997 PhD
- Stanford 1991~1994 MS
- Google, IBM, Microsoft

### Timeline generated

#### **Columbia University:** 1997~2020

found evidence:

<e1> Luis Gravano </e1> is a Professor of Computer Science at <e2> Columbia University </e2>.

<e1> Luis Gravano </e1> is a professor in the Computer Science department at <e2> Columbia University </e2> - see what their students are saying about them or leave a rating.

<e1> Luis Gravano </e1> | <e2> Columbia University </e2> | United States of America.

#### **Stanford University:** 1995~1997

found evidence:

<e1> Luis Gravano </e1> received the PhD degree in computer science from <e2> Stanford University </e2> in 1997.

#### **IBM:** 1991~1992

New Institutions found by Web scraping:

##### **Computer Science Department**

<e1> Luis Gravano </e1> has been on the faculty of the <e2> Computer Science Department </e2>, Columbia.

<e1> Luis Gravano </e1> is Principal Investigator in <e2> Computer Science Department </e2> with research interest databases maintainance.

##### **Columbia**

<e1> Luis Gravano </e1> joined <e2> Columbia </e2> in 1997.

##### **Google**

At <e1> Google </e1>, <e2> Luis Gravano </e2> was a Senior Research Scientist in 2001.

**Evaluation:** Mostly accurate besides failing to predict **Microsoft**. Notice that there is multiple evidence that augment the fact that Luis Gravano was indeed affiliated with **Columbia University** and **Stanford University**. Also, web scraping allowed us to catch **Google**, which wasn't here at first hand with the base openalex timeline.

## Abdussalam Alawini

- University of Illinois Urbana-Champaign 2018~
- University of Pennsylvania 2016~2018 PostDoc
- Portland State 2011~2016 PhD
- Portland State 2011 MS

### Timeline generated

**University of Illinois Urbana-Champaign:** 2018~2021

**University of Pennsylvania:** 2017~2018

found evidence:

<e1> Abdussalam Alawini </e1> ALAWINI, Postdoctoral Fellow | Cited by 100 | of <e2> University of Pennsylvania </e2>, PA (UP) | Read 14 publications | Contact Abdussalam ALAWINI.

**Portland State University:** 2013~2015

New Institutions found by Web scraping:

**University Of Illinois**

<e1> Abdussalam Alawini </e1> is a professor in the Computer Science department at <e2> University Of Illinois </e2> at Urbana - Champaign - see what their students are saying.

**UP**

<e1> Abdussalam Alawini </e1> University of Pennsylvania | <e2> UP </e2> - ResearchGate.

**University Of Illinois at**

<e1> Abdussalam Alawini </e1> at <e2> University Of Illinois at </e2> Urbana - Champaign.

**University of Illinois**

<e1> Abdussalam Alawini </e1>'s research works | <e2> University of Illinois </e2>, Urbana.

**University of Illinois at**

<e1> Abdussalam Alawini </e1> is a Teaching Assistant Professor in the Department of Computer Science at the <e2> University of Illinois at </e2> Urbana-Champaign.

<e1> Abdussalam Alawini </e1> · Teaching Assistant Professor - <e2> University of Illinois at </e2> Urbana-Champaign · Report · Report · Experience · Education · Volunteer Experience.

**Evaluation:** Successfully predicted every institution and mostly correct with the years the author was affiliated with each institution. The evidence generated by RBERT model is somewhat messy as NER model didn't recognize "Urbana-Champaign" as a part of the name of institution.

## Nan Jiang

- University of Michigan 2011~2017
- Microsoft 2017~2018
- UIUC 2018~

### Timeline generated

**University of Illinois Urbana-Champaign:** 2016~2021

**University of Michigan–Ann Arbor:** 2014~2016

New Institutions found by Web scraping:

**University of Illinois at Urban ##a**

<e1> Nan Jiang </e1> - Assistant Professor - <e2> University of Illinois at Urban ##a </e2>.

**Missouri University of Science and Technology**

<e1> Nan Jiang </e1> was an Assistant Professor at <e2> Missouri University of Science and Technology </e2> (Missouri S&T) from 2016 to 2020 and a Postdoctoral Associate Researcher in.

**The Ohio State University**

I am <e1> Nan Jiang </e1>-Jiang Jiang, a fifth year PhD candidate in Linguistics at <e2> The Ohio State University </e2> working on computational linguistics.

**Evaluation:** Predicted every institution and years besides **Microsoft**. It is highly suspicious that the institutions found by web scraping are affiliated with Professor Nan Jian based on the context of evidence. The evidence suggested with **Ohio State University** is not relevant to Professor Nan Jiang; it should be someone else with the same name. Therefore, we need our model to classify if the given sentence(evidence) really belongs to the author of our focus.

## Kevin Chen-Chuan Chang

- National Taiwan University
- Stanford University
- University of Illinois Urbana-Champaign

### Timeline generated

**University of Illinois Urbana-Champaign:** 2013~2022

**Stanford University:** 1998~1998

New Institutions found by Web scraping:

**University Of Illinois at Urban ##a**

<e1> Kevin Chen-Chuan Chang </e1> Chang at <e2> University Of Illinois at Urban ##a </e2> - Rate My Professors.

**National Taiwan University**

<e1> Kevin Chen-Chuan Chang </e1> received a BS from <e2> National Taiwan University </e2> and PhD from.

**University of Illinois**

<e1> Kevin Chen-Chuan Chang </e1>'s research works | <e2> University of Illinois </e2>.

**University Of Illinois**

<e1> Kevin Chen-Chuan Chang </e1> Chang is a professor in the Computer Science department at <e2> University Of Illinois </e2> at Urbana - Champaign - see what their students are saying about.

**University of Illinois at**

<e1> Kevin Chen-Chuan Chang </e1> is a Professor in Computer Science, <e2> University of Illinois at </e2> Urbana-Champaign.

<e1> Kevin Chen-Chuan Chang </e1> is a Professor in Computer Science, <e2> University of Illinois at </e2> Urbana-Champaign, where he leads the FORWARD Data Lab.

<e1> Kevin Chen-Chuan Chang </e1> is a professor with the <e2> University of Illinois at </e2> Urbana-Champaign.

Bio: <e1> Kevin Chen-Chuan Chang </e1> is an academic researcher from <e2> University of Illinois at </e2> Urbana-Champaign.

**Evaluation:** Successfully predicted all the institutions besides the fact that there are some years corresponding to UIUC missing. Here, we could realize the power of web scraping and RBERT model as it successfully found out the author's educational background that was missing in Openalex, such as **National Taiwan University** in this case.

## Jonathan Aldrich

- University of Washington: 2001~2003
- Carnegie Mellon University: 2003~

### Timeline generated

#### **Carnegie Mellon University:** 2003~2022

found evidence:

<e1> Jonathan Aldrich </e1> - Professor - <e2> Carnegie Mellon University </e2> -

LinkedIn.

#### **University of Washington:** 2001~2003

New Institutions found by Web scraping:

##### **School of**

Author Bio: <e1> Jonathan Aldrich </e1> is a professor in the <e2> School of </e2>.

**Evaluation:** Perfectly predicted the institutions and their corresponding years

## Peter Bailis

- Sisu Data: 2018~
- Stanford University: 2016~2020
- MIT: 2015~2016
- UC Berkeley: 2011~2015
- Harvard College: 2007~2011

### Timeline generated

**Sisu Data:** 2021~2022

**Stanford University:** 2015~2021

**University of California, Berkeley:** 2010~2015

New Institutions found by Web scraping:

#### **Si ##su**

Deciphering Your Data with <e1> Peter Bailis </e1>, Founder and CEO of <e2> Si ##su </e2>.

<e1> Peter Bailis </e1> is the co-founder and CEO of <e2> Si ##su </e2> which is a software that empowers people to make better decisions with data.

#### **National Science Foundation**

<e1> Peter Bailis </e1> wrote his PhD thesis on large-scale data management and is a recipient of a <e2> National Science Foundation </e2> Graduate Research Fellowship.

#### **Si ##su Data**

<e1> Peter Bailis </e1> is the founder and CEO of <e2> Si ##su Data </e2>, the fastest and most comprehensive diagnostic platform for structured data.

<e1> Peter Bailis </e1> is the founder and CEO of <e2> Si ##su Data </e2>, a data analytics platform that helps users understand the key drivers behind critical busi.

**Evaluation:** Predicted every institution besides **MIT** and **Harvard College**. As mentioned above, our current web scraping (evidence finding) process works great for some authors, but there is definitely a space to improve. I believe we can try more advanced web scraping instead of relying on Trafilatura

## Sarah Chasins

- UC Berkeley: 2012~
- Swarthmore College: 2012

### Timeline generated

**University of California, Berkeley:** 2020~2021

**Swarthmore College:** 2011~2011

New Institutions found by Web scraping:

#### **UC Berkeley**

<e1> Sarah Chasins </e1> is an Assistant Professor at <e2> UC Berkeley </e2>.

#### **E ##EC ##S**

<e1> Sarah Chasins </e1> - Assistant Professor of <e2> E ##EC ##S </e2> - LinkedIn.

#### **University of California**

<e1> Sarah Chasins </e1> · Assistant Professor of EECS at <e2> University of California </e2>, Berkeley · Report · Report · About · Experience · Education · Honors & Awards · Websites.

**Evaluation:** Predicted every institution and year, but again, some years missing for **UC Berkeley**. Information extraction from selected paragraphs might be useful to fill in this gap.

## Armando Fox

- UC Berkeley 2006~
- Stanford University 1998~2007
- UC Berkeley 1994~1998
- UIUC 1992~1994

### Timeline generated

**University of California, Berkeley:** 1995~2021

**Stanford University:** 1998~2007

**University of Illinois Urbana-Champaign:** 1992~1993

New Institutions found by Web scraping:

#### **Digital Learning Strategy Equity**

<e1> Armando Fox </e1> is Professor of Computer Science, Faculty Advisor for <e2> Digital Learning Strategy Equity </e2> Advisor, and Computer Science Diversity Officer.

Electrical Engineering & Computer Science Department

<e1> Armando Fox </e1> is a Professor in Berkeley's <e2> Electrical Engineering & Computer Science Department </e2> as well as the Faculty Advisor for Digital Learning Strategy.

#### **Associated Students of Stanford University**

During his previous time at Stanford,<e1> Armando Fox </e1>received teaching and mentoring awards from the <e2> Associated Students of Stanford University </e2>, the Society of Women.

#### **University of California Berkeley**

<e1> Armando Fox </e1> at <e2> University of California Berkeley </e2> - Rate My Professors.

<e1> Armando Fox </e1> is a professor in the Computer Science department at <e2> University of California Berkeley </e2> - see what their students are saying about them or leave.

**Evaluation:** Timeline generator worked great here. Predicted every institution and years. It even found some of the smaller bodies of some schools, such as **Associated Students of Stanford University**.



## Jennifer Listgarten

- UC Berkeley 2018~
- Microsoft 2007~2017
- University of Toronto

### Timeline generated

**University of California, Berkeley:** 2018~2022

**Microsoft:** 2006~2018

**University of Toronto:** 2003~2006

**Cross Cancer Institute:** 2003~2003

New Institutions found by Web scraping:

#### **Department of Electrical Engineering and Computer Science**

<e1> Jennifer Listgarten </e1> is a Professor in the <e2> Department of Electrical Engineering and Computer Science </e2>, and Center for Computational Biology.

**Evaluation:** Perfectly predicted every institution and most of the years are accurate as well

## Sanmi Koyejo

- University of Illinois Urbana-Champaign
- Stanford University
- University of Texas at Austin

### Timeline generated

**University of Illinois Urbana-Champaign: 2018~2021**

New Institutions found by Web scraping:

**University Of Illinois at Urban ##a**

<e1> Sanmi Koyejo </e1> at <e2> University Of Illinois at Urban ##a </e2> - Rate My Professors.

**University Of Illinois**

<e1> Sanmi Koyejo </e1> is a professor in the Computer Science department at <e2> University Of Illinois </e2> at Urbana - Champaign - see what their students are saying about.

**Evaluation:** Only predicted UIUC. This might be due to the limitation of scraped data provided by Trafilatura. While evaluating the timeline generator, I consistently felt the lack of knowledge caused by Trafilatura, and therefore, I would like to try different approaches and web scraping tools to collect the data via web scraping.