DEPARTMENT RANKINGS

H. V. Jagadish University of Michigan

2016: State of Affairs

Better known CS Dept rankings in US

- US News and World Report
- National Research Council (1995, 2010)

2018: CS Rankings Got Worse

Before

NO DATA and no transparency

US News & World Report Rankings of CS Departments

NOW

BAD DATA and no transparency

US News & World Report Rankings of Global Universities in CS

2018: CS Rankings Got Worse

Before

NO

US N

Best Global Universities for Computer Science







NOW

BAI

US N

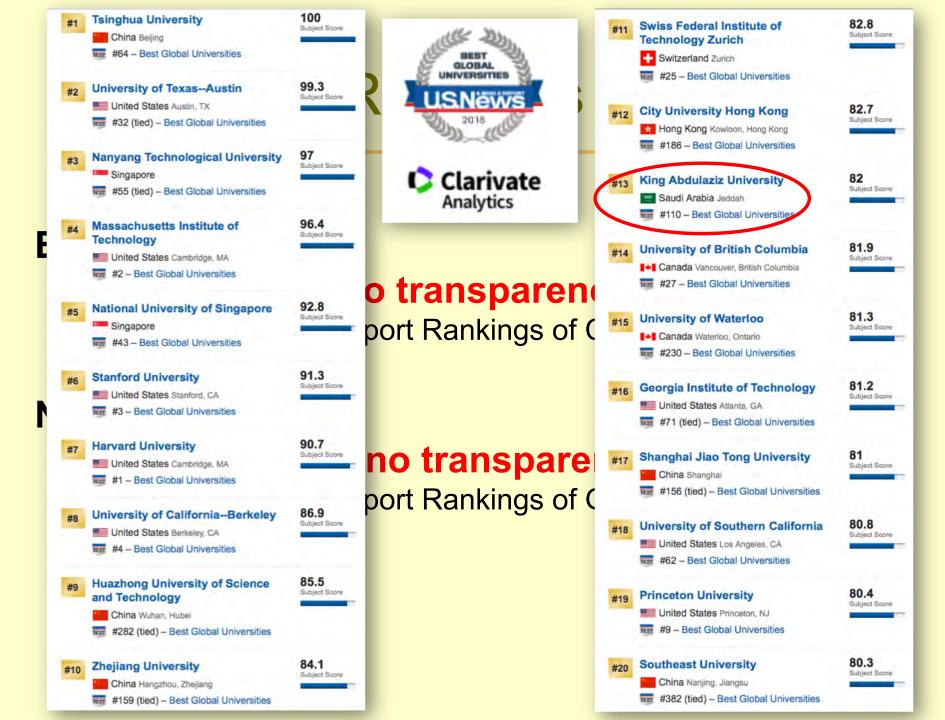
Computer science is the study of computers and their uses, and the field comprises a wide range of subjects. They include software engineering and design, artificial intelligence theory, programming languages, information systems and information technology. These are the world's top universities for computer science, which were ranked based on their reputation and research in the field.







in CS



Uses & abuses of rankings

Efficient way to inform decisions

- Choosing a PhD program (especially foreign students)
- Applying for an academic position (PhD graduates)
 Imposes structure on the field
- e.g., CRA salary comparisons of "like" institutions.

Used in discussions between departments and administration

- Rewards for ratings improvements
- Funding for remedial action when ratings fall
- Reality check on claims

Problems with Ranking Schemes

- Trailing indicator
- Imposes a value system
 - Different people have different needs and will flourish in different environments

Nb: Horror vacui -Parmenides 485BC

= "Nature abhors a vacuum"

Infeasible to decide:

There should be no ranking system.

Two Years Ago at Snowbird

Department Rankers and Rankings: Truths and Consequences

H. V. Jagadish, U. Michigan Fred Schneider, Cornell U. Steve Furber, U. Manchester Bob Morse, US News & WR

DEPARTMENT RANKING COMMITTEE

CURRENT

Steve Blackburn (ANU, Australia)

Emery Berger (UMass, US)

Carla Brodley (Northeastern, US)

H. V Jagadish, Chair (Michigan, US)

Kathryn McKinley (Google US)

Mario Nascimento (Alberta, Canada)

Minjeong Shin (ANU, Australia)

Lexing Xie (ANU, Australia)

Andy Bernat, Ex-officio (CRA, US)

CHARGE in 2016

- Determine if there is a data gathering role for CRA to assist with the various program ranking efforts.
- 2. Determine a set of metrics that are collectable or obtainable from others and that we would be comfortable standing behind.
- 3. Work with the various program ranking efforts to determine which of these metrics would be of value to them and which they would commit to using.

Work with Rankers

- US News was the major partner identified.
- Initial conversations were very positive.
- Ultimately, they decided they want to stick with reputation-based rankings for US programs, at least for now.

Rankings: can't live with'em, can't live without'em

The latest US News and World Report (USN&WR) ranking of Computer Science (CS) at global universities does a grave disservice to USN&WR readers and to CS departments all over the world [...] We urge the community to ignore the **USN&WR** rankings of Computer



Science
CRA Statement on US News and World Report Rankings of Computer Science Universities November 2017

GOTO Principles Adopted by CRA

Evaluation methodologies must be datadriven and meet at least the following criteria:

- Good data
 - data have been cleaned and curated
- Open
 - o data available, regarding attributes measured, at least for verification
- Transparent
 - process and methodologies are entirely transparent
- Objective
 - based on measurable attributes

Is There Hope?

• Yes! gotorankings.org





- Several efforts supporting GOTO principles
- Today, we will hear about three prominent efforts

Three Distinguished Panelists

Kuansan Wang

- Managing Director, Microsoft Research Outreach
- Microsoft Academic Services (http://aka.ms/msracad)

Kathryn S McKinley

- Senior Scientist, Google
- Chair, CRA Committee on Metrics
- csmetrics.org

Emery Berger

- Professor, UMass Amherst
- csrankings.org

Microsoft Academic Services

Kuansan Wang, Microsoft

Microsoft Academic Services

- How to empower researchers to do more and achieve more, with AI
- System components:
 - Knowledge acquisition
 - Machine reading all web documents
 - Microsoft Academic Graph (MAG)
 - Knowledge reasoning and inference
 - Search & Recommendation
 - academic.microsoft.com
 - Reinforcement learning
 - Predicting citation behaviors as crowd-sourced impact assessments

Microsoft Academic Graph

Annual growth rate 9.6% (x2@7.25 years) ~2M new publications/month Bi-weekly updates available upon request!



Publications

175,369,863

Coming soon



Fields of Study

229,031

Learn more



Conferences

4,028

Learn more



Authors

211,481,059

Learn more



Journals

47,975

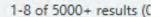
Learn more



25,309

Learn more





Date Range

2009 V to 201



- Geoffrey E. Hinton
- Ilya Sutskever
- Yoshua Bengio
- Ross B. Girshick
- Jian Sun

Show more >

Affiliation

- Microsoft
- Google
- Stanford Univer
- University of Ca
- University of To
- University of Ox
- Massachusetts Institute of Technology

Authors known for this topic

Karen Simonyan (University of Oxford), Andrew Zisserman (University of Oxford)

Abstract: In this work we investigate network depth on its accuracy in the

Sort by: Relevance

Networks for Large-

Semantic, not just keyword matching

* Citations (12,986) * 🕒 Download 🕏 Share 🏴 Cite

ImageNet Classification with Deep

Canadational Named Naturorks

Top institutions for the topic

15, pp 1097-1105

Sutskever (University of Toronto),

neural network to classify n the ImageNet LSVRC-



Default Ranking in MA

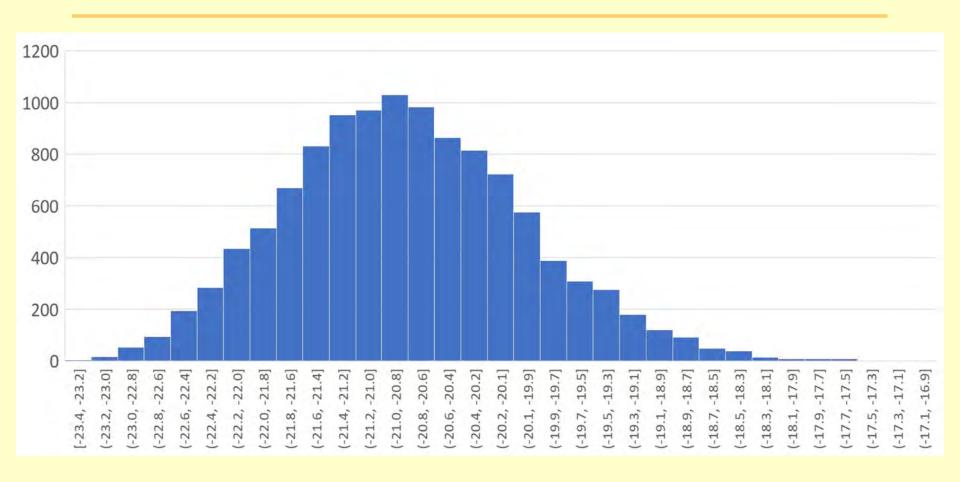
Lessons:

- "The Google Scholar experiment" (2014)
- "Promise and Pitfalls of Extending Google's PageRank Algorithm to Citation Networks" (2008)

Saliency:

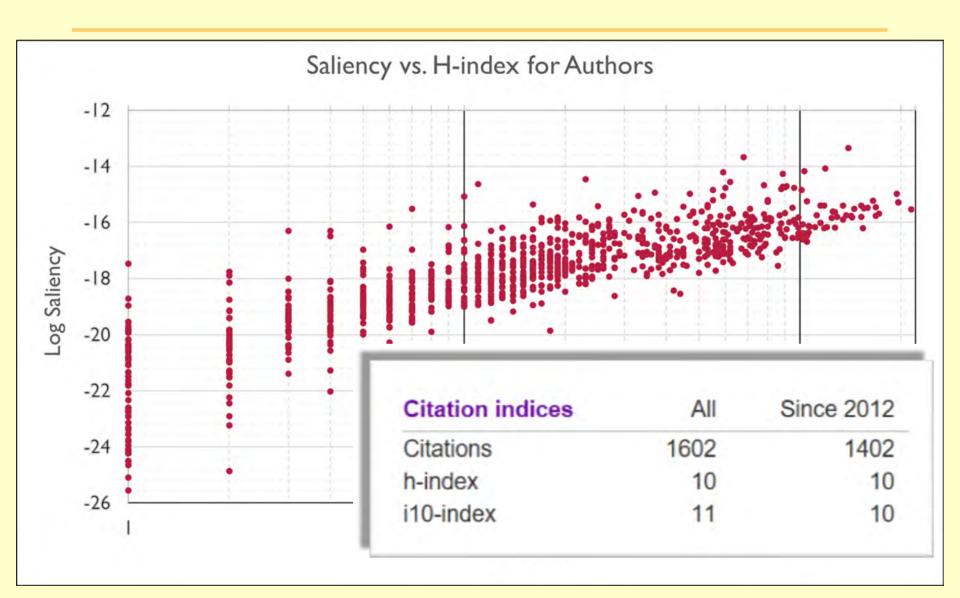
- Probability of being referred to by other salient entities, aging over time
- Compute for every type of nodes on the graph
- Reinforcement learning for latent parameters

Paper saliencies in a venue

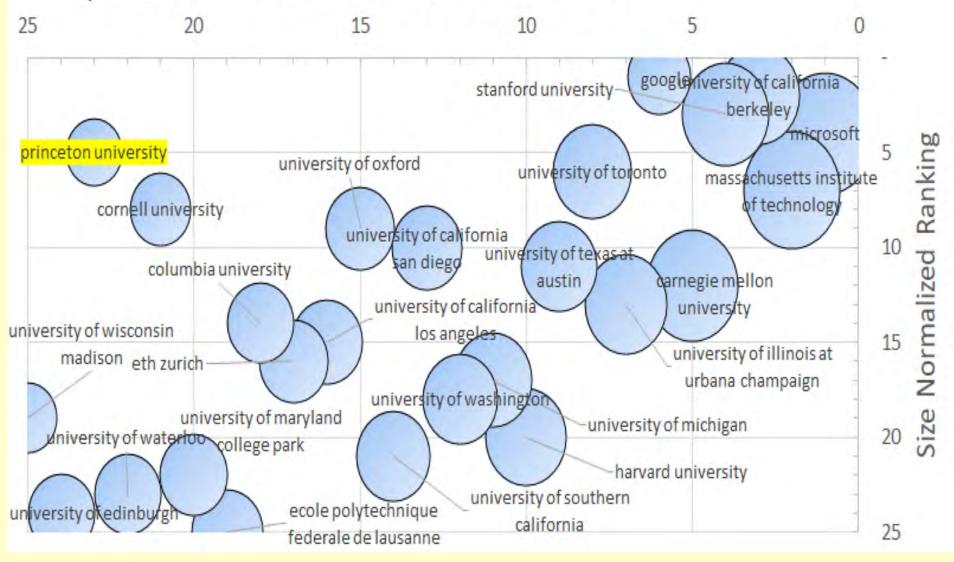


Publication venue: not a good predictor of impact

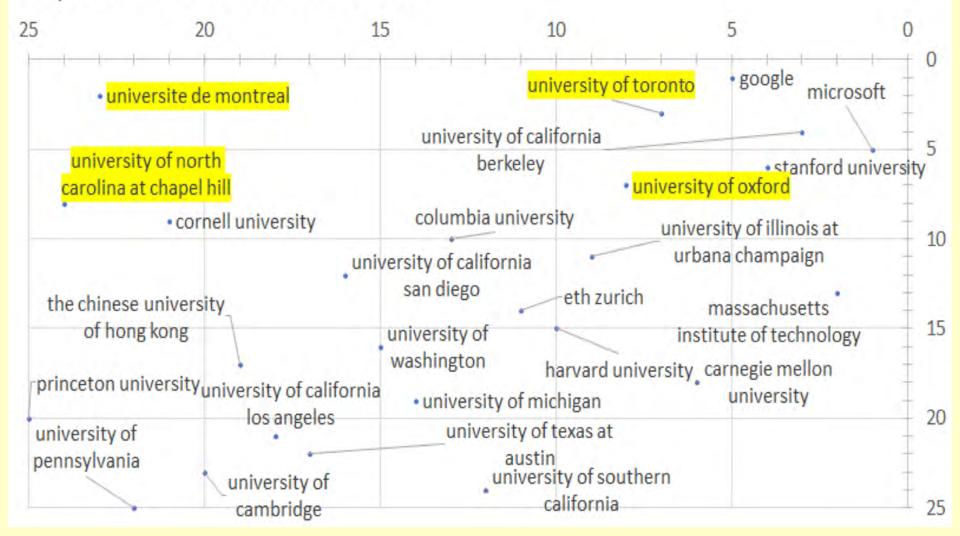
Productivity+impact both captured



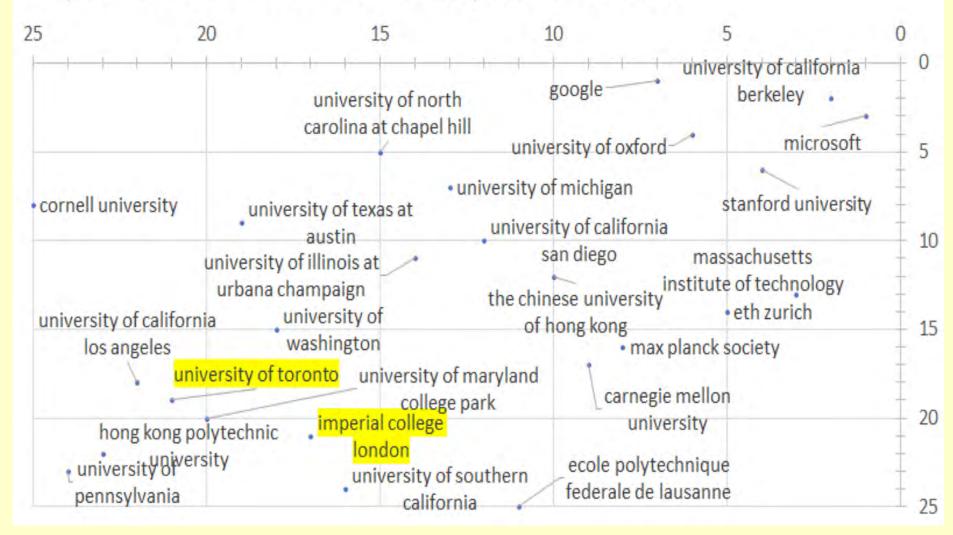
Top 25 Research institutions in CS



Top 25 institutions in Al



Top 25 institutions in Computer Vision



Open data to tell nuanced stories are available

MAS Supports GOTO

- Web scale reading enables cross-validation
 - Include preprints, conf proceedings, patents,...
 - CVs, homepages for author disambiguation
- MAG: Openly available upon request
 - https://aka.ms/msracad
- Promote open source ranking algorithms
 - Check out our GitHub repository
- "...unreasonable to expect departments halfway around the world will have anything close to an accurate assessment of each other" - CRA statement Nov. 2017

csmetrics.org

Kathryn S McKinley, Google

csmetrics.org 2017

Institutional ranking

Audience

- University administrators
- Faculty, students, parents

Key features

GOTO methodology

Configurable data

Configurable publication metrics

- Citations for past
- Counts for future

now 20 r entries		$\alpha = 0.3 0.0$ Combined = (Measured+z) ^{0.7} x (Predicted+z) ^{0.3}		
Rank	Institution CRA Academic •	Measured	Predicted	Combined
1	Carnegie Mellon University	101698	9993	50702
2	Massachusetts Institute of Technology	96060	9600	48135
3	Stanford University	99710	8200	47126
4	University of California, Berkeley	89053	7094	41689
5	University of Illinois at Urbana-Champaign	76028	5791	35116
6	University of Washington	59375	5097	28426
7	Georgia Institute of Technology	49079	5547	25518
8	University of Texas at Austin	48915	4807	24388
9	University of Toronto	48211	3822	22537
10	University of Michigan	41134	5391	22358
11	University of California, San Diego	44335	3557	20799
12	Cornell University	39143	4330	20221
13	Princeton University	40781	3663	19791
14	University of Maryland, College Park	37293	3895	18936
15	University of Southern California	35742	3837	18299
16	University of California, Los Angeles	34138	3257	16870
17	University of Wisconsin- Madison	30573	3808	16366
18	Columbia University	29094	3715	15691
19	Purdue University	28306	3396	14983
20	University of Pennsylvania	27351	2547	13418

csmetrics.org 2017

Institutional ranking

Audience

- University administrators
- Faculty, students, parents

Key features

GOTO methodology

Configurable data

Configurable publication metrics

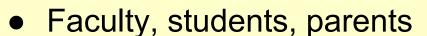
- Citations for past
- Counts for future

100	entries		a = 0.3 0.0 Combined =	(Measured+c)0.7 x	(Pred/cted+t) ^{U,2}	
Rank	Institution CRA Academic ▼		Measured	Predicted	Combined +	
1	Carnegie Mellon University		101698	9993	50702	
2	Massachusetts Institute of Technology		96060	9600	48135	
3	Stanford University		99710	8200	47126	
4	University of California, Berkeley	21	Pennsylvania State University			
5	University of Illinois at Urbana-Champaign	22	University of Massachusetts Amherst			
6	University of Washingto	23	University of Waterloo			
7	Georgia Institute of Technology	24	University of California, Irvine			
8	University of Texas at A	25	Univer	sity of Minne	sota	
9	University of Toronto				oota	
10	University of Michigan	26	Harvard University			
11	University of California, Diego	27	University of British Columbia			
12	Cornell University	28	Ohio S	Ohio State University		
13	Princeton University	29	Arizona	Arizona State University		
14	University of Maryland, College Park	30	Rutger	Rutgers University		
15	University of Southern California	31	New Y	ork Universit	у	
16	University of California, Angeles	32	Univers Barbar	sity of Califor a	rnia, Santa	
17	University of Wisconsin- Madison	33	Johns Hopkins University			
18	Columbia University	2.4	University of North Carolina a		Carolina at	
19	Purdue University	34	Chapel Hill			
20	University of Pennsylva	35	Duke I	Jniversity		

Ranking Institutions vs Individuals

Ranking institutions

- University administrators
 - Activity in area X
 - O Do we need to invest more?
 - Did prior investments pay off?
 - Collaborations



- Where should we go?
- Activity in area X
- For graduate students, choosing faculty mentor(s)



Institution = Publications of *all* affiliated authors

Curated & configurable by venue Criteria: rigorous peer reviewed venues Data

- Started with CRA venue list
- 209 conference venues
- 80 Journal venues
- Publications: cleaned DBLP data
- Citations from MAG
- 6646 Institutions, cleaned with MAG

Publication Data

Philosophy

- Rigorous peer review
- Inclusive: impact & activity beyond top tier
- Configurable

Example

Programming Languages



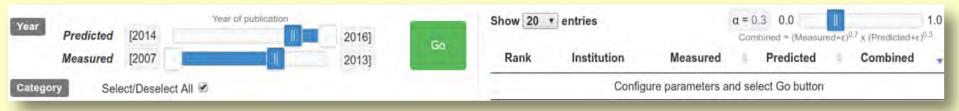
Cleaning DBLP publication data

Full research papers from 2007-2016 Download with DBLP filters Add new filters for titles, front matter, etc. Number of papers per year verification Reviewed by hand all years for errors ACM & IEEE sources have errors List to MAG for citations

Areas

Year		Year of publication	
Predicted	[2014	2016]	Go
Measured	[2007	2013]	
Category	lect/Deselect	All 🗷	
ALGORITHMS ARC	HITECTURE	ARTIFICIAL INTELLIGENCE BIOINFORMATICS	CIRCUITS
CRYPTO	GRAPHY DATA	A MINING DATABASES DESIGN AUTOMATIO	N
DISTR	RIBUTED SYSTE	EMS EMBEDDED & REAL TIME GRAPHICS	
HIGH PE	RFORMANCE C	OMPUTING HUMAN COMPUTER INTERACTION	N
INFORMAT	ION RETRIEVAL	MACHINE LEARNING MOBILE MULTIMEE	OIA
NATURAL I	ANGUAGE PRO	OCESSING NETWORKS OPERATING SYSTE	MS
PARALLEL COM	IPUTING PER	FORMANCE ANALYSIS PROGRAMMING LANG	UAGES
ROBOTICS SECUR	RITY SOFTWA	ARE ENGINEERING SYSTEMS THEORY VE	RIFICATION
	VISION	WEB TECHNOLOGIES OTHER	
Venue Weight	Geo Mean	Predicted = Σ Paper Count x Venue Weight	
	All	▼ Select Venue Type	
Venue List	Туре	Fullname	Weight
⊘ 3DIM	С	International Conference on 3D Digital Imaging and Modeling	2

Metrics



Measured citations for older papers

Paper = 1 divided by authors

Each institution receives fractional credit, never changes

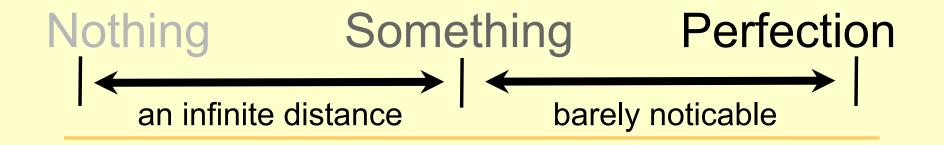
Predicted weighted counts of recent papers

Weight recent by venue impact (configurable)

Venue impact = geomean of citations per paper

Configuration

area, venue, venue weighting, year, past, predicted weighting of past/predicted



Better, but not perfect

- DORA Declaration of Research Assessment says do not count by venue impact!
- Citations practices change, differ by area
- Citation gaming
- Faculty size, Faculty current institution
- Lack of Interdisciplinary coverage (e.g. no Nature)
- Missing other metrics, e.g., funding, awards, etc.
- Volunteer workforce

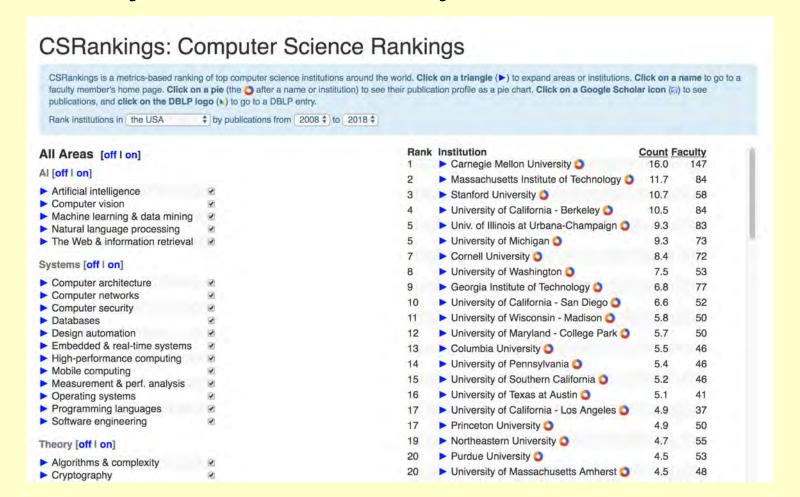
GOTO: code+data on github https://github.com/csmetrics/csmetrics.org

FAQ and User guide: http://csmetrics.org/faq/
https://github.com/csmetrics/csmetrics.org#quickstart

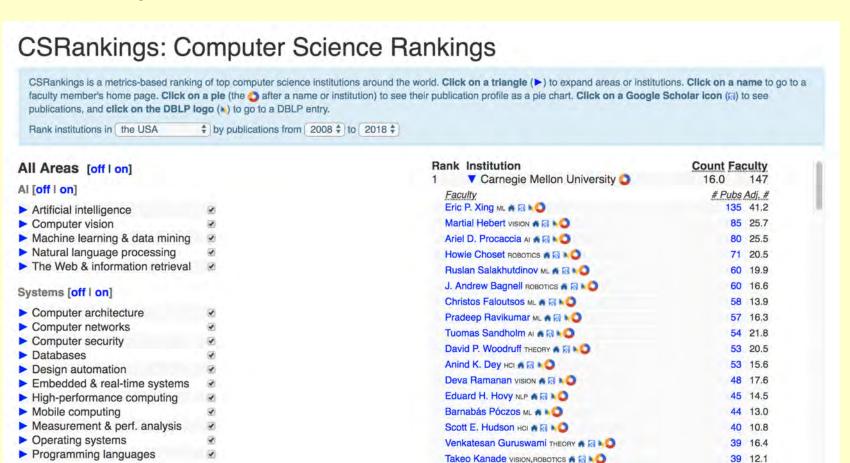
csrankings.org

Emery Berger
University of Massachusetts,
Amherst

Publicly announced July 2016



Faculty-centric, conference pub-focused



Lorrie Faith Cranor Hol € № № William W. Cohen NLP € № №

37 11.9

Software engineering

Theory [off | on]

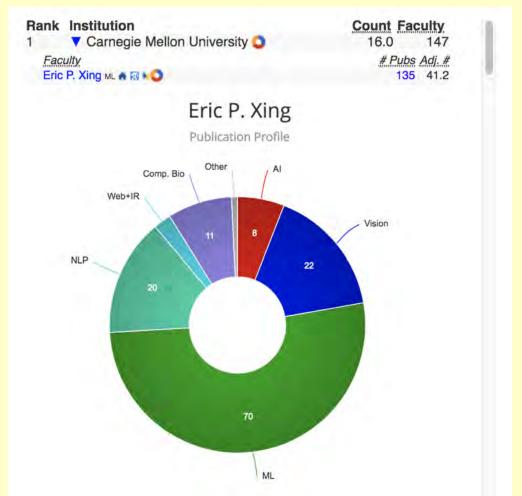
Faculty-centric, conference pub-focused

Rank Institution	Count Faculty
1 V Carnegie Mellon University	16.0 147
Faculty	# Pubs Adj. #
Eric P. Xing ML # ₪ MO	135 41.2
Martial Hebert vision # 🕅 📢	85 25.7
Ariel D. Procaccia AI A A A	80 25.5
Howie Choset ROBOTICS A N NO	71 20.5
Ruslan Salakhutdinov ML 🕷 🔯 🔌 🔘	60 19.9
J. Andrew Bagnell новотіся 🛊 🔞 📢	60 16.6
Christos Faloutsos ML # R NO	58 13.9
Pradeep Ravikumar ML 🛊 🔞 📢	57 16.3
Tuomas Sandholm AI # 🗐 🕦	54 21.8
David P. Woodruff THEORY # 🖓 📢	53 20.5
Anind K. Dey нсі 🛦 🖟 🔌	53 15.6
Deva Ramanan vision 🛊 🖟 😂	48 17.6
Eduard H. Hovy NLP # 🕅 💆	45 14.5
Barnabás Póczos ML # NO	44 13.0
Scott E. Hudson но 🛊 🖟 📞	40 10.8
Venkatesan Guruswami тнеоку 🚜 🐼 👡	39 16.4
Takeo Kanade vision, ROBOTICS # 🕅 📢	39 12.1
Lorrie Faith Cranor но 🛊 🗟 📢	39 6.3
William W. Cohen NLP ★ 🕅 🔌	37 11.9
Anupam Gupta THEORY # 17 NO	37 11.1
Vipul Goyal CRYPTO, THEORY A 🐼 📢	34 12.7
Tom M. Mitchell NLP # [3] *	34 11.3
David G. Andersen NETWORKS # 🕅 📢	33 6.9

dept =
current faculty

count
normalized
across
disciplines
(normalizes for
publication rate)

Faculty-centric, conference pub-focused



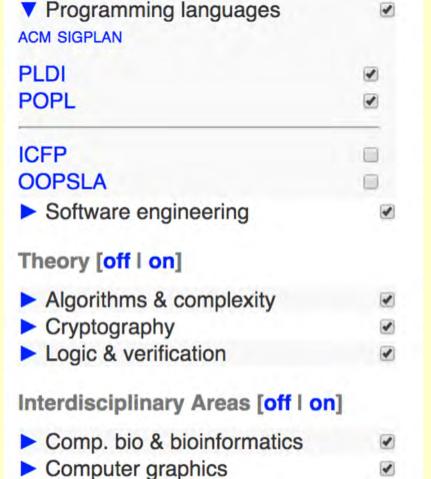
Counts papers in top conferences

All Areas [off on]	
Al [off on]	
Artificial intelligence	-
Computer vision	1
▶ Machine learning & data mining	4
Natural language processing	1
► The Web & information retrieval	•
Systems [off on]	
Computer architecture	1
Computer networks	1
Computer security	1
Databases	1
Design automation	1
► Embedded & real-time systems	1
▶ High-performance computing	1
▶ Mobile computing	1
► Measurement & perf. analysis	1
➤ Operating systems	1
► Programming languages	4
➤ Software engineering	1

Counts papers in top conferences

Programming languages	
Software engineering	
Theory [off I on]	
Algorithms & complexity	
Cryptography	•
Logic & verification	
Interdisciplinary Areas [off on]	
Comp. bio & bioinformatics	•
Computer graphics	
Economics & computation	
Human-computer interaction	•
➤ Robotics	
Visualization	

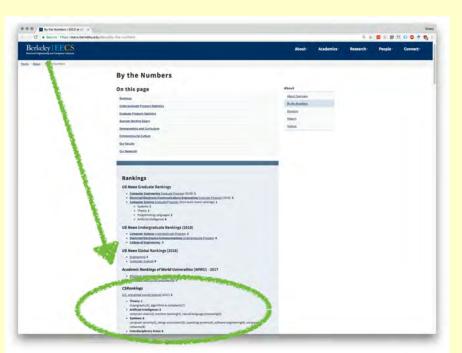
Counts papers in top conferences



+ optional below-the-fold conferences

Publicly announced July 2016

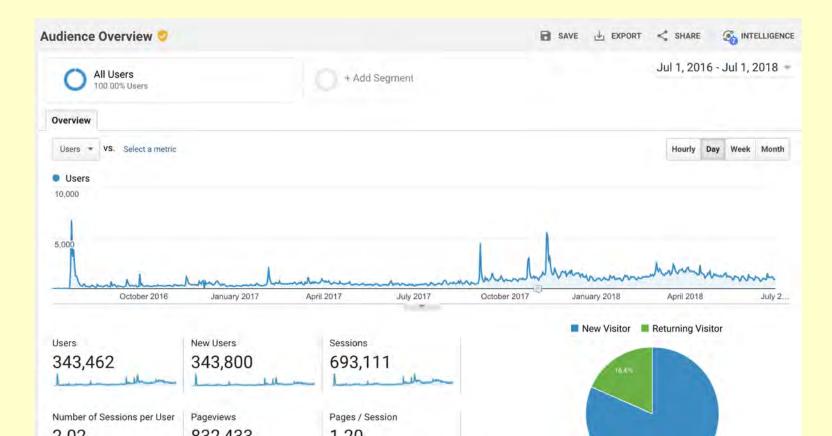
 Now cited by growing list of department web pages (Berkeley, Michigan, Edinburgh, Cornell, CMU...)





Publicly announced July 2016

> 300K users to date



Can select specific subfields of interest (with "permalinks")

CSRankings: Computer Science Rankings

CSRankings is a metrics-based ranking of top computer science institutions around the world. Click on a triangle (▶) to expand areas or institutions. Click on a name to go to a faculty member's home page. Click on a pie (the after a name or institution) to see their publication profile as a pie chart. Click on a Google Scholar icon (⋈) to see publications, and click on the DBLP logo (★) to go to a DBLP entry.

Rank institutions in the USA \$\dip \text{by publications from 2008 \$\dip \text{to 2018 \$\dip}

All Areas [off on]		Rank Institution		Count Fa	aculty
		1	Carnegie Mellon University 🔾	56.5	80
Al [off on]		2	Cornell University	30.3	40
Artificial intelligence	②	3	➤ Stanford University <a>○	26.9	32
Computer vision	Ø	4	► Georgia Institute of Technology	22.0	31
 Machine learning & data mining Natural language processing 	Ø	5	► University of Maryland - College Park ②	19.1	26
The Web & information retrieval	•	6	➤ University of California - Berkeley 🔾	19.0	40
Systems [off on]		7	► Massachusetts Institute of Technology ○	18.7	48
		8	➤ University of Michigan <	18.4	33
Computer architecture		9	► University of Massachusetts Amherst 🥥	16.3	29
Computer networks		10	► Univ. of Illinois at Urbana-Champaign •	15.9	32
 Computer security Databases 	in .	11	➤ University of Pennsylvania Ο	15.3	26
Design automation		12	► University of Southern California ②	14.7	26
► Embedded & real-time systems		13	► University of Texas at Austin •	14.5	17
▶ High-performance computing		13	► University of Washington ○	14.5	30
► Mobile computing	0	15	Columbia University	12.8	23
Measurement & perf. analysis		10	- Odiumbia University	12.0	20



Yann LeCun shared Charles Sutton's post.

57 mins · 🚱

Charles Sutton writes about CS department rankings.

As we know, the US News & World Report ranking is ridiculous and should be ignored.

CSrankings.org has considerably better methodology, and allows you to see how CS departments stack up in subfields of computer science.

For example, NYU is quite good in ML, vision, NLP, theory, verification, crypto, graphics and visualization. Click these subfields and NYU is number 6, behind CMU, MIT, Stanford, Berkeley and Cornell:

http://csrankings.org/...

With all subfields turned on, NYU is 19th, largely because it's a relatively small department with a few areas of excellence and entire areas with little or no presence.

The most important question to pick a place for your PhD is "who do you want to work with?"

"who do you want to work with?"

CSRankings: Computer Science Rankings

CSRankings is a metrics-based ranking of top computer science institutions around the world. Click on a triangle (>) to expand areas or institutions. Click on a name to go to a faculty member's home page. Click on a pie (the 💍 after a name or institution) to see their publication profile as a pie chart. Click on a Google Scholar icon 🔞 to see publications, and click on the DBLP logo (*) to go to a DBLP entry.

3

\$ by publications from 2008 \$ to 2018 \$ Rank institutions in the USA

All Areas [off | on]

Al [off | on]

- Artificial intelligence
- Computer vision Machine learning & data mining
- Natural language processing
- ➤ The Web & information retrieval

Systems [off | on]

- Computer architecture
- Computer networks
- Computer security
- Databases
- Design automation
- ► Embedded & real-time systems High-performance computing
- Mobile computing Measurement & perf. analysis
- Operating systems
- Programming languages

Rank Institution Count Faculty Massachusetts Institute of Technology <a>O 14.4 16 University of California - Berkeley 13 8.3

7.8

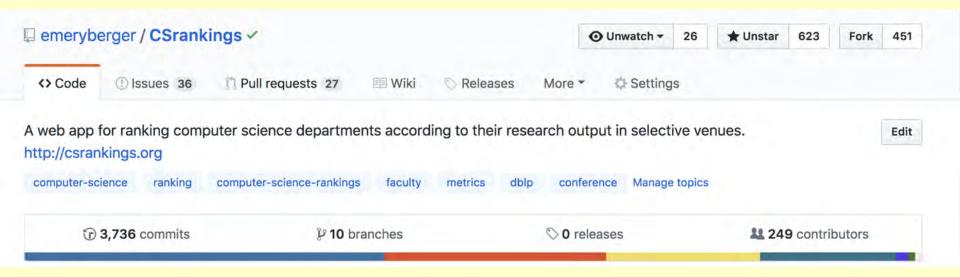
12

University of Michigan <a>O

- # Pubs Adj. # Faculty Jason Flinn os, ARCH, MOBILE # 13 10 2.0 Manos Kapritsos os # R NO 0.9 Baris Kasikci os A 🔞 🍆 0.8 Peter M. Chen ARCH # 17 NO. 0.7
- Mosharaf Chowdhury NETWORKS # [3] 1 0.5 0.5 Zhuoqing Morley Mao MOBILE, METRICS, SECURITY A NO Scott A. Mahlke ARCH A R NO 0.4
- Michael J. Cafarella DB # 13 100 0.3 Satish Narayanasamy ARCH # 13 10 0.3 0.2
- Stéphane Lafortune 🙀 🕅 📢 Harsha V. Madhyastha NETWORKS # [3] NO 0.2 Thomas F. Wenisch ARCH # 13 NO 0.2
- Stanford University 6.7 10

GOTO - all code & data on GitHub

https://github.com/emeryberger/CSrankings



FAQ here: http://csrankings.org/faq.html

What Do We Want?

- Reasonable people can disagree about precisely what to measure and how to combine measures into a rank.
- Some may even prefer to have a multidimensional score rather than a strict linear rank ordering.
- Let us focus today on some principles we would like to see followed.

GOTO Principles

Evaluation methodologies must be datadriven and meet at least the following criteria:

- Good data
 - data have been cleaned and curated
- Open
 - o data available, regarding attributes measured, at least for verification
- Transparent
 - process and methodologies are entirely transparent
- Objective
 - based on measurable attributes

Discussion Period

Straw Poll 1

All metrics and rankings of Computing departments should follow GOTO principles.

Straw Poll 2

The CRA should fund a project to promote GOTO rankings.

Straw Poll 3

I will personally ignore any ranking that does not follow GOTO principles.