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 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 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

#13 King Abdulaziz University
Saudi Arabia, Jeddah
#110 – Best Global Universities
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)
1. 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.
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
Evaluation methodologies must be data-driven and meet at least the following criteria:

- **Good data**
  - data have been cleaned and curated

- **Open**
  - 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!
Authors known for this topic
Semantic, not just keyword matching
Top institutions for the topic
Default Ranking in MA

• Lessons:
  – “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

princeton university

cornell university

university of oxford

university of california san diego

university of toronto

university of texas at austin

carnegie mellon university

university of illinois at urbana champaign

university of michigan

harvard university

university of southern california

ecole polytechnique federale de lausanne

university of edinburgh

university of maryland college park

university of washington

microsoft

massachusetts institute of technology

stanford university

google

university of california berkeley

university of wisconsin madison

eth zurich

university of waterloo
Top 25 institutions in AI

universite de montreal
university of north carolina at chapel hill
cornell university
university of toronto
google
microsoft
stanford university
university of oxford
the chinese university of hong kong
princeton university
university of california los angeles
university of pennsylvania
university of cambridge
university of washington
university of california san diego
university of california berkeley
columbia university
university of illinois at urbana champaign
eth zurich
massachusetts instiute of technology
harvard university, carnegie mellon university
university of michigan
university of texas at austin
university of southern california
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
Institutional ranking

**Audience**
- University administrators
- Faculty, students, parents

**Key features**

- GOTO methodology
- Configurable data
- Configurable publication metrics
  - Citations for past
  - Counts for future
Institutional ranking

Audience
- University administrators
- Faculty, students, parents

Key features
- GOTO methodology
- Configurable data
- Configurable publication metrics
  - Citations for past
  - Counts for future
Ranking Institutions vs Individuals

Ranking institutions
● University administrators
  ○ Activity in area X
  ○ Do we need to invest more?
  ○ Did prior investments pay off?
  ○ Collaborations

● Faculty, students, parents
  ○ 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
**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
Nothing | Something | Perfection
| an infinite distance | | barely noticeable |

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](https://github.com/csmetrics/csmetrics.org)

csrankings

Publicly announced July 2016
csrankings

Faculty-centric, conference pub-focused
Faculty-centric, conference pub-focused
dep = current faculty
count normalized across disciplines (normalizes for publication rate)
csrankings

Faculty-centric, conference pub-focused
Counts papers in *top* conferences
csrankings

Counts papers in *top* conferences
csrankings

Counts papers in *top* conferences

- Programming languages
  - ACM SIGPLAN
    - PLDI
    - POPL
  - ICFP
    - OOPSLA
- Software engineering
- Theory [off | on]
  - Algorithms & complexity
  - Cryptography
  - Logic & verification
- Interdisciplinary Areas [off | on]
  - Comp. bio & bioinformatics
  - Computer graphics

+ 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 “permalink”)
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?”
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 multi-dimensional 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 data-driven and meet at least the following criteria:

- **Good data**
  - data have been cleaned and curated

- **Open**
  - 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.