ACM HPDC 2019, Phoenix, Arizona, USA

Perspectives on High-Performance Computing in a Big Data World – Part A

The 28th International Symposium on High-Performance Parallel and Distributed Computing

Geoffrey Fox | June 27, 2019

gcf@indiana.edu, http://www.dsc.soic.indiana.edu/, http://spidal.org/

Outline

Supported by National Science Foundation through Awards
- 1443054 CIF21 DIBBs: Middleware and High Performance Analytics Libraries for Scalable Data Science
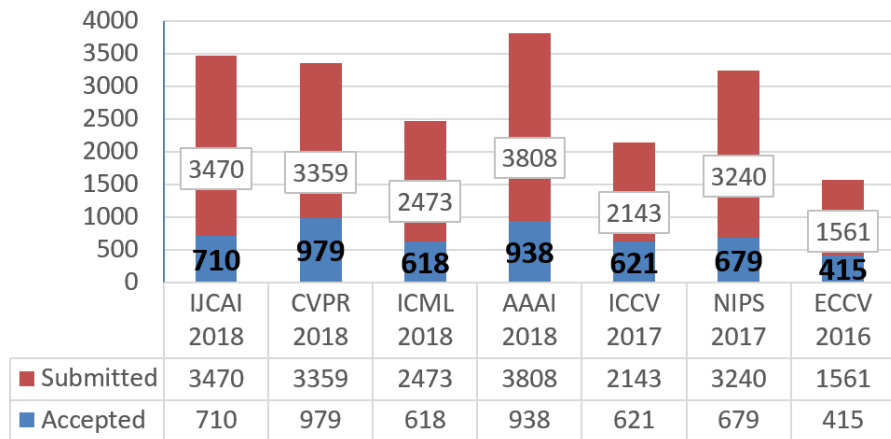- 1720625 Network for Computational Nanotechnology - Engineered nanoBIO Node

- Discussion of Communities from HPC to Edge to Big Data to Machine Learning and Cloud
- Aligning with Industry in an AI First world
- "Machine Learning for Systems" or MLforHPC
  - Scenarios
  - Examples
  - Some open questions
- Conclusions

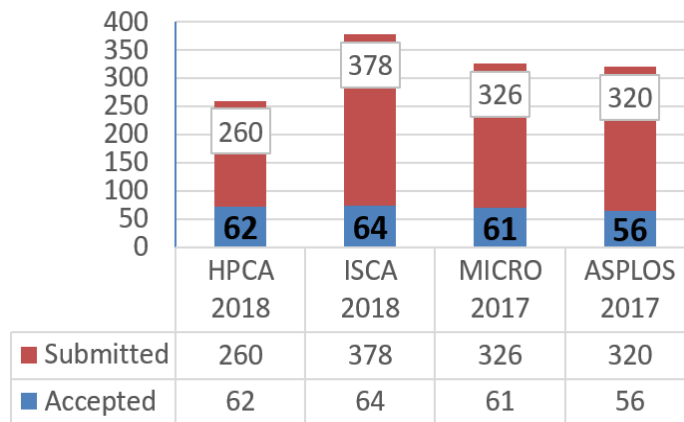# Data on the Evolution of Interests and Communities

Qian Depei: The difference between AI and computer architecture academic research is hot

Artificial intelligence main conference paper statistics

	IJCAI 2018	CVPR 2018	ICML 2018	AAAI 2018	ICCV 2017	NIPS 2017	ECCV 2016
Submitted	3470	3359	2473	3808	2143	3240	1561
Accepted	710	979	618	938	621	679	415

■ Submitted   ■ Accepted

H5index   45   158   91   56   89   101   98

Architecture main conference paper statistics

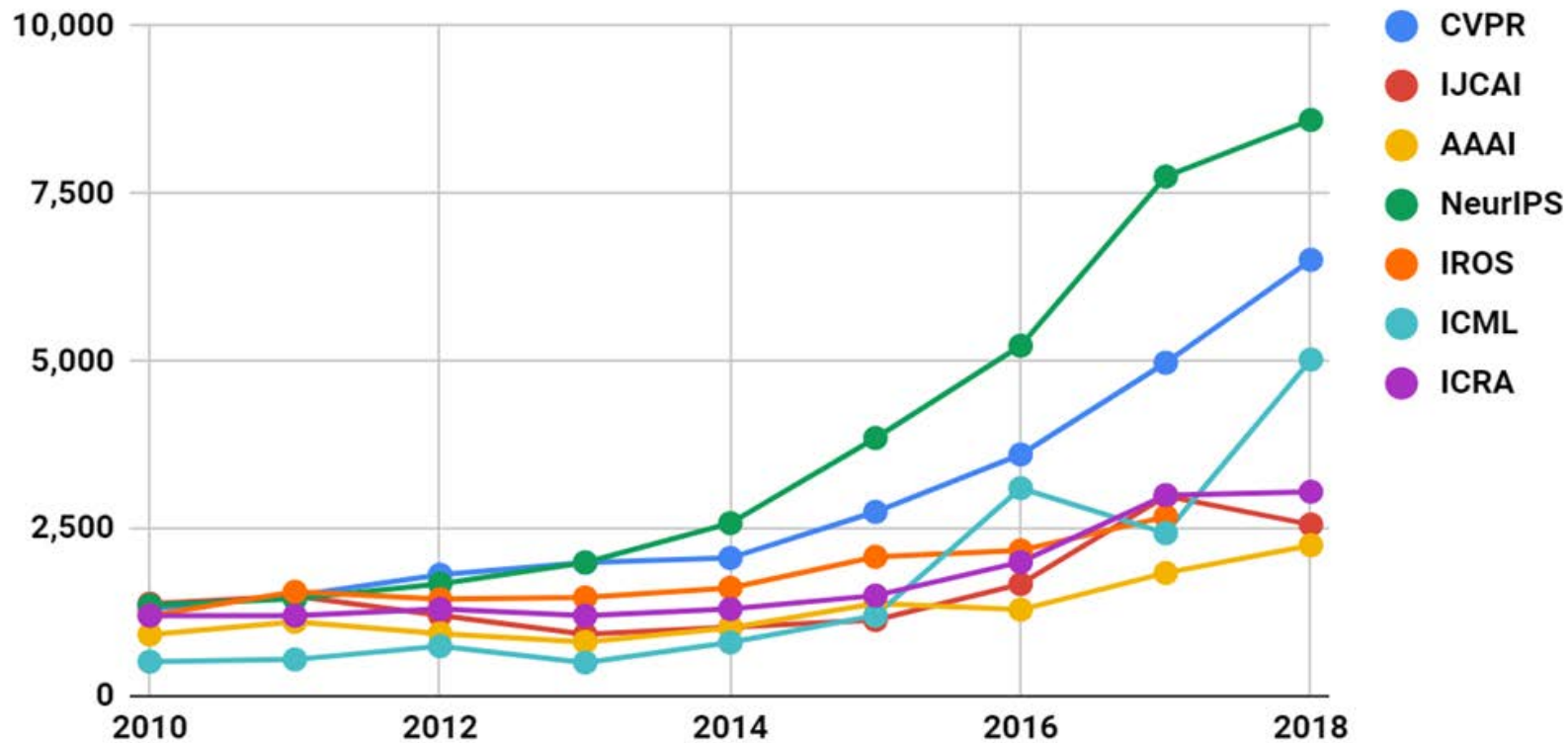|	HPCA 2018	ISCA 2018	MICRO 2017	ASPLOS 2017
Submitted	260	378	326	320
Accepted	62	64	61	56

■ Submitted   ■ Accepted

46   54   41   50

Number of papers published 4960 vs 243

Papers Submitted: Comparing 4 Conference Types

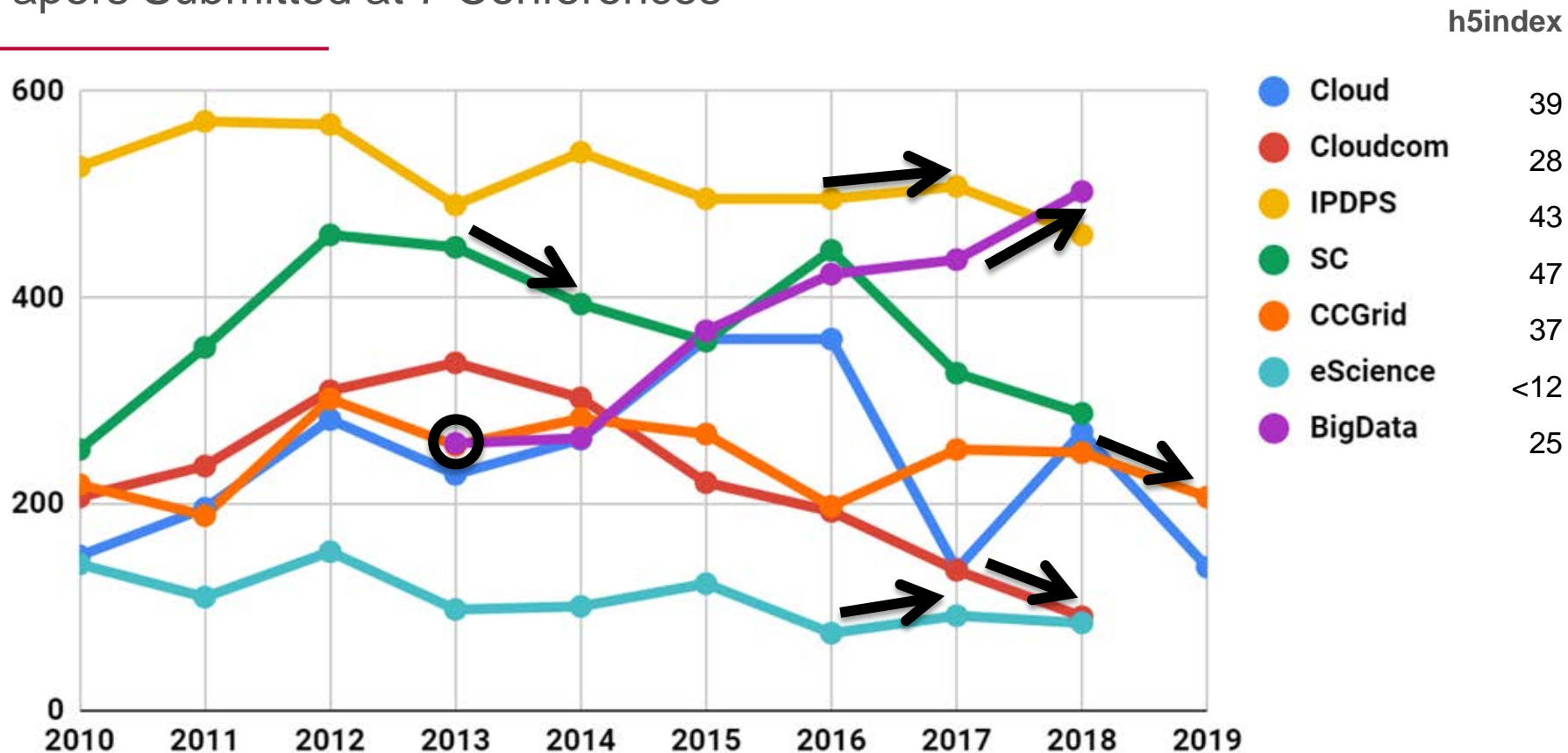SumCI: SC, eScience, CCGrid, IPDPS

SumCloud: IEEE Cloud, Cloudcom
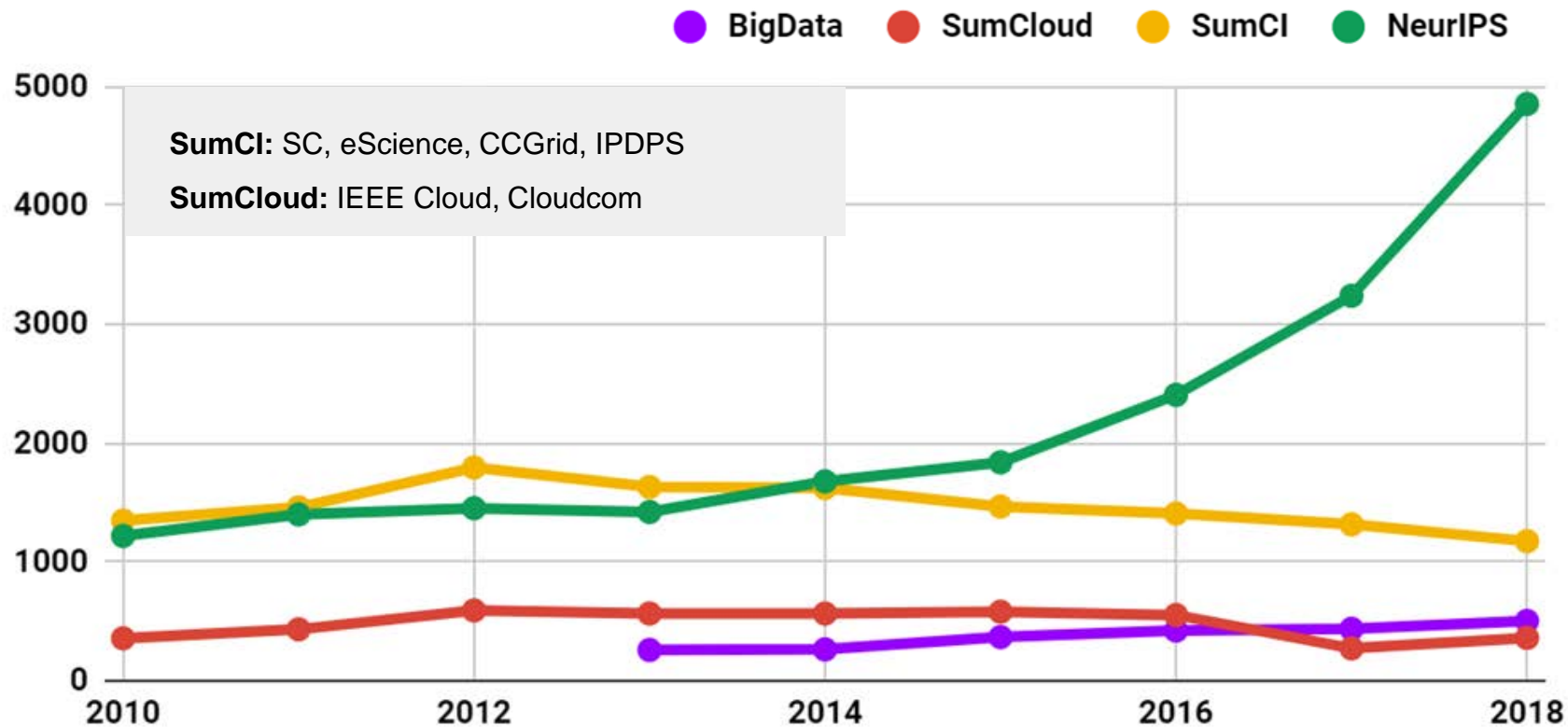
Attendance at Major AI Conferences

Papers Submitted: Comparing 7 Conferences and 2 Sums

HPDC not included as talk given there SoCC (other major cloud conference) not included as could find no statistics

Legend: Cloud, Cloudcom, IPDPS, SC, CCGrid, eScience, BigData, SumCloud, SumCI, NeurIPS
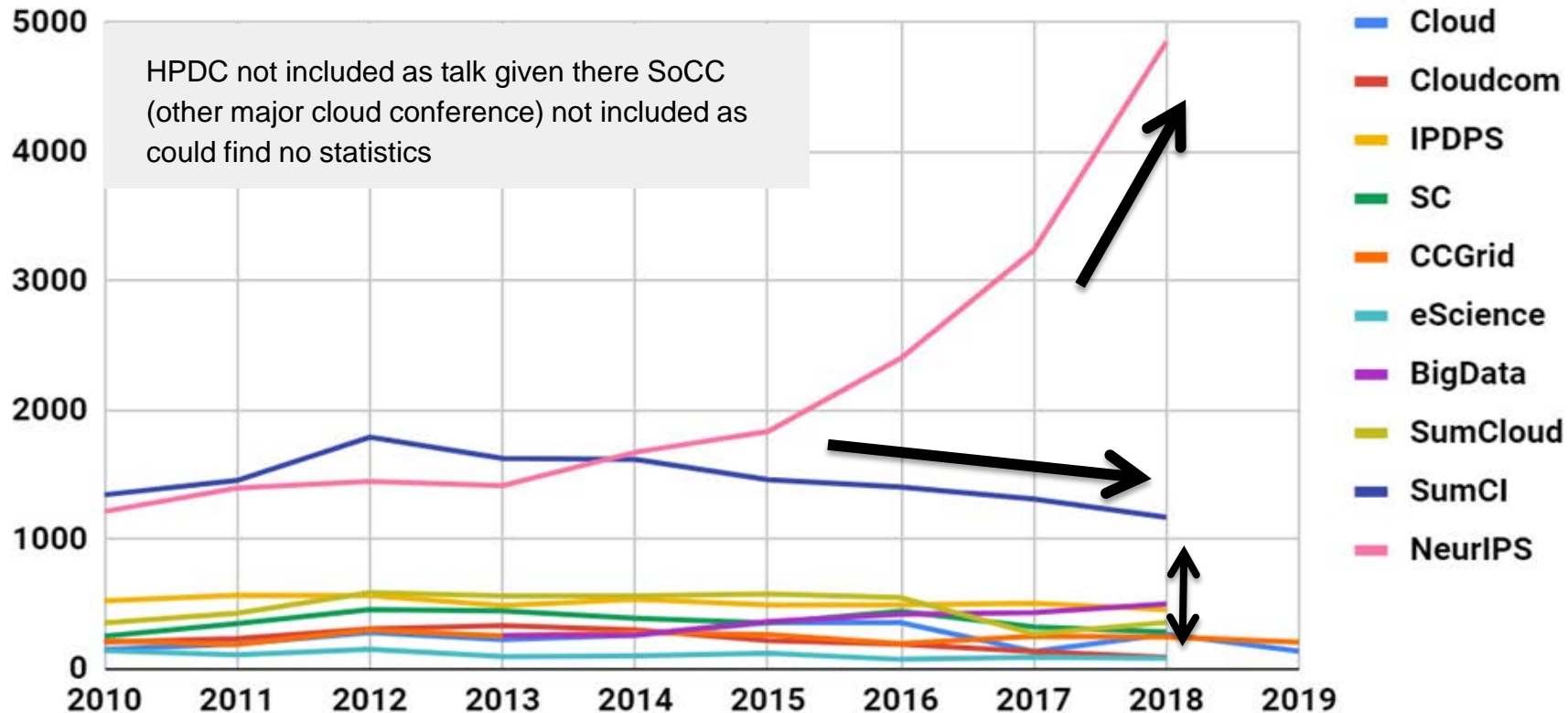
Papers Submitted at 7 Conferences

h5index

Conference	h5index
Cloud	39
Cloudcom	28
IPDPS	43
SC	47
CCGrid	37
eScience	<12
BigData	25

Arxiv Publications from Aiindex.org

In 2017, absolute number of papers are

AI: 23,922   **CS:** 383,279
~~**All:** 3,032,731~~
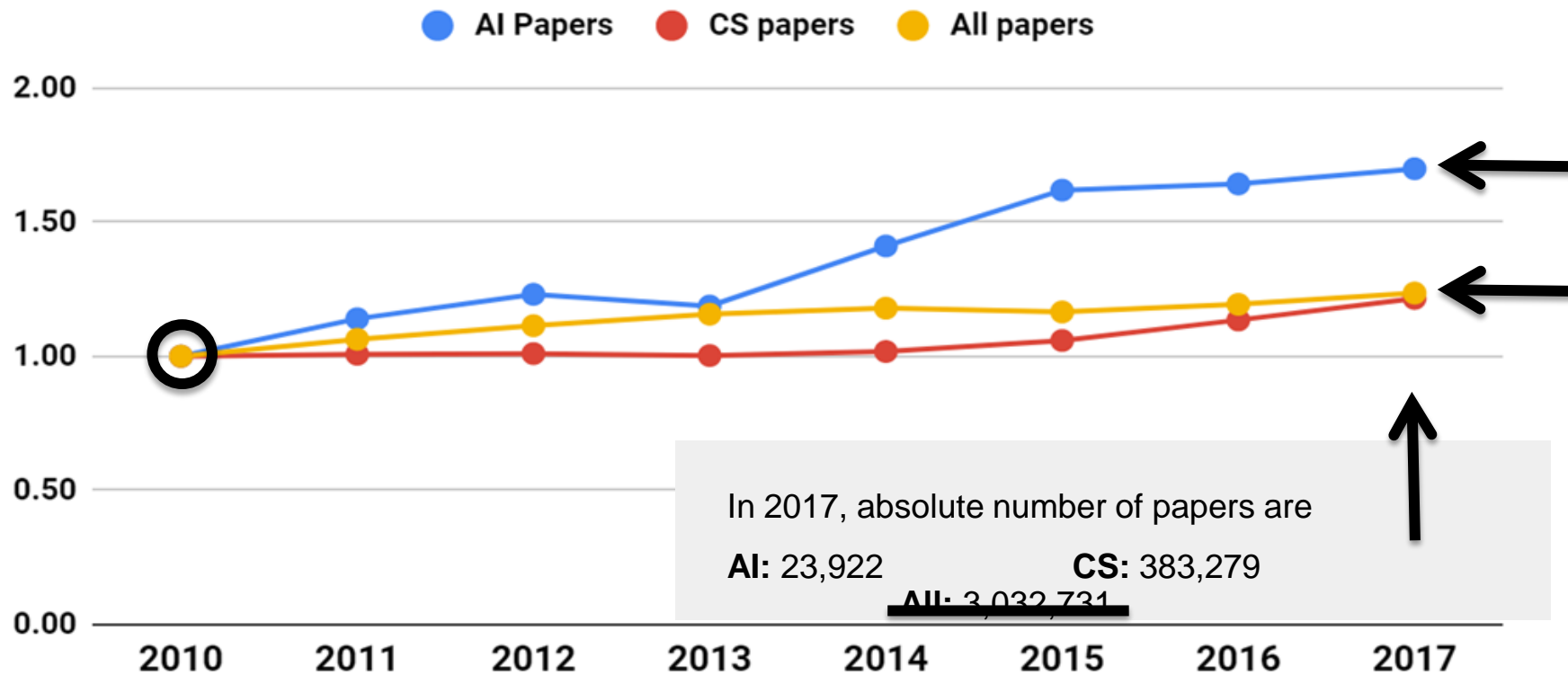
H5-index correlated with conference size
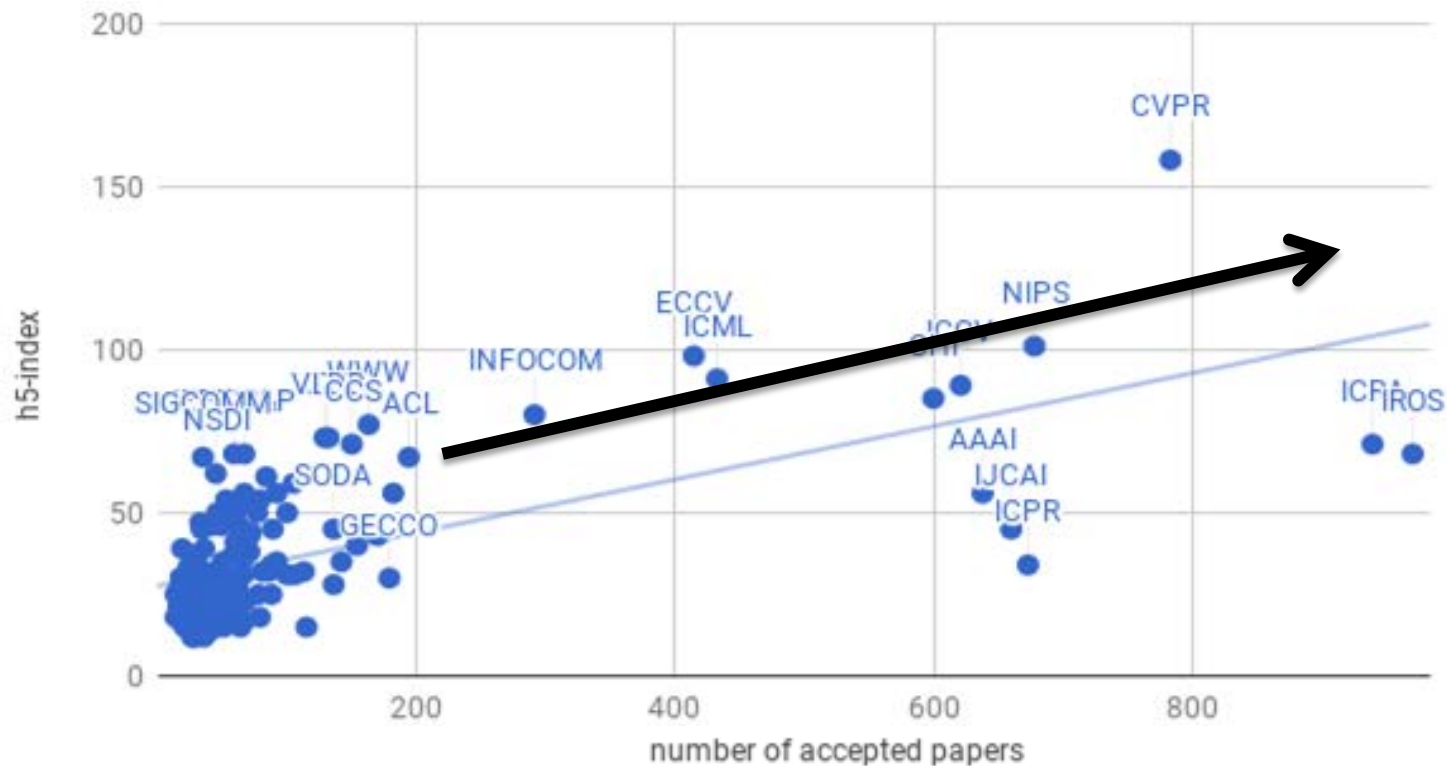
H5 is h-index calculated over the last 5 years produced in July 2018

H5-index Conferences: AI, Big Data, Cloud, Systems, Other

158 CVPR: Conference on Computer Vision and Pattern Recognition
101 NeurIPS: Neural Information Processing Systems
98 ECCV: European Conference on Computer Vision
91 ICML: International Conference on Machine Learning
89 ICCV: International Conference on Computer Vision
85 CHI: Computer Human Interaction
80 INFOCOM: Joint Conference of the Computer and Communications Societies
77 WWW: International World Wide Web Conferences
73 VLDB: International Conference on Very Large Databases
73 SIGKDD: International Conference on Knowledge discovery and data mining
71 ICRA: International Conference on Robotics and Automation
56 AAAI: Assoc. Adv. AI Conference on Artificial Intelligence
54 ISCA: International Symposium on Computer Architecture
50 IROS: International Conference on Intelligent Robots and Systems
50 ASPLOS: International Conference on Architectural Support for Programming Languages and Operating Systems
47 SC: International Conference on High Performance Computing, Networking, Storage and Analysis
46 HPCA: International Symposium on High Performance Computer Architecture
45 IJCAI: International Joint Conference on Artificial Intelligence
43 BMVC: British Machine Vision Conference
43 IPDPS: International Symposium on Parallel & Distributed Processing
41 MICRO: International Symposium on Microarchitecture

39 CLOUD: International Conference on Cloud Computing
39 OSDI: Symposium on Operating Systems Design and Implementation
37 OOPSLA: SIGPLAN International Conference on Object-Oriented Programming, Systems, Languages, and Applications
37 PPOPP: SIGPLAN Symposium on Principles & Practice of Parallel Programming
37 CCGrid: International Symposium on Cluster Computing and the Grid
34 ICIP: International Conference on Image Processing
34 ICPR: International Conference on Pattern Recognition
30 SoCC: Symposium on Cloud Computing
29 ECAI: European Conference on Artificial Intelligence
29 HPDC: International Symposium on High Performance Distributed Computing
28 CloudCom: International Conference on Cloud Computing Technology and Science
26 ICS: International Conference on Supercomputing
25 Big Data: International Conference on Big Data
21 CLUSTER: International Conference on Cluster Computing
21 SPAA: Symposium on Parallelism in Algorithms and Architectures
20 ICPP: International Conference on Parallel Processing
18 ICCSA: International Conference on Computational Science and Its Applications
**15 SBAC-PAD:** International Symposium on Computer Architecture and High Performance Computing
14 DS-RT: International Symposium on Distributed Simulation and Real-Time Applications

Some didn't make h5-index cut of >=12

Some large areas: Google Trends last 5 years *(Topics unless otherwise stated)*

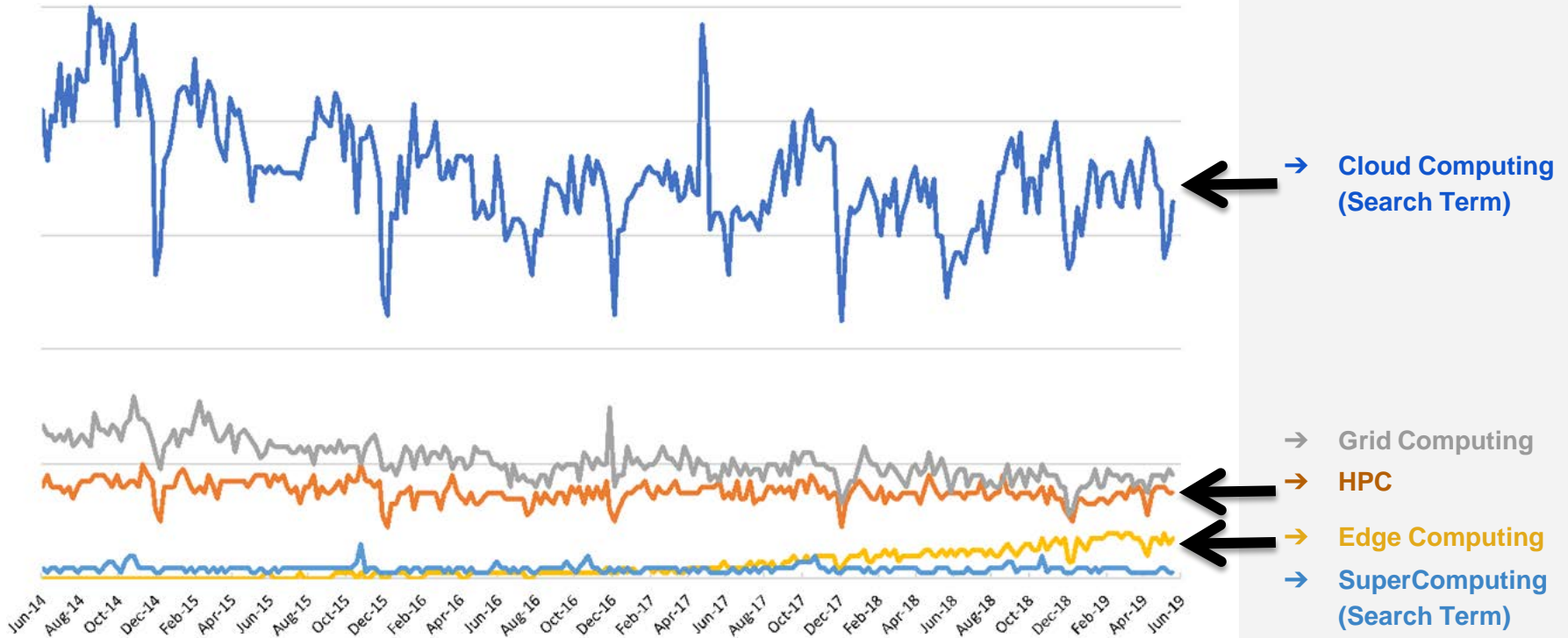**Search terms require exact match - topics are broader but sometimes are not available**
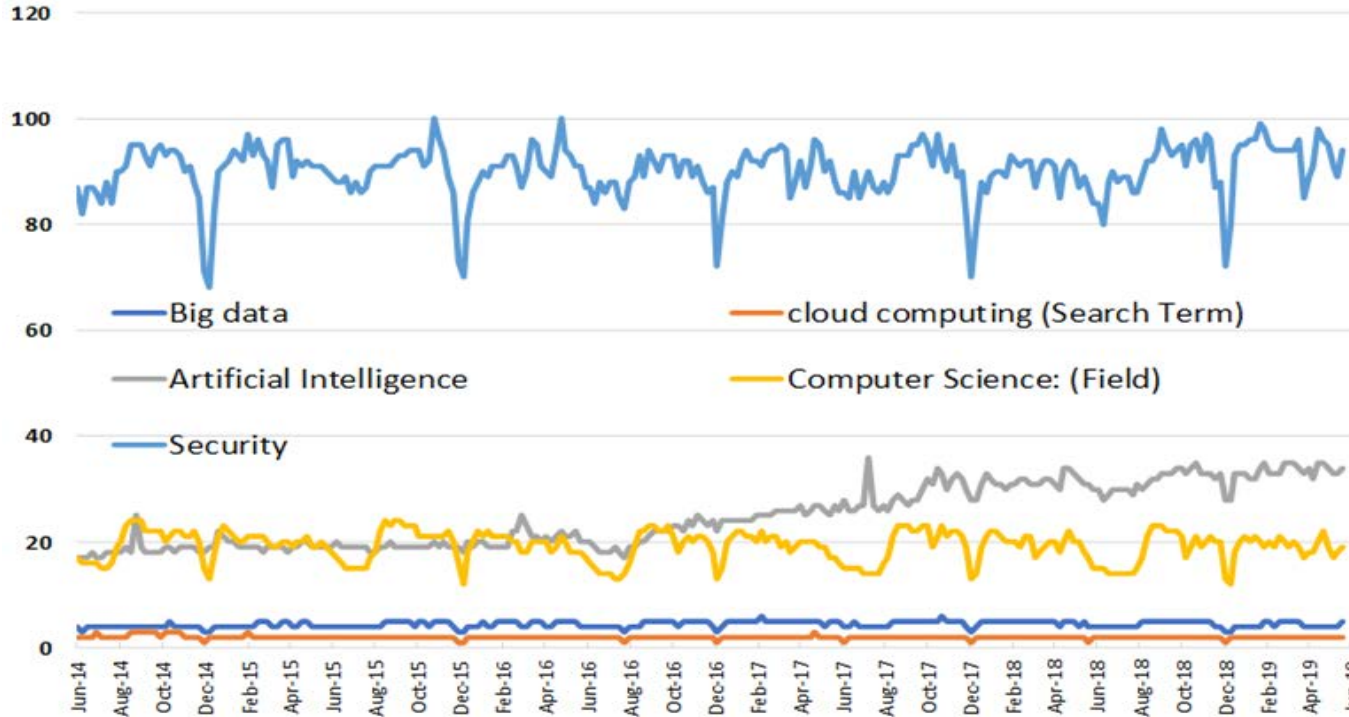
→ **Security**

→ **Artificial Intelligence**

→ **Computer Science (Field)**

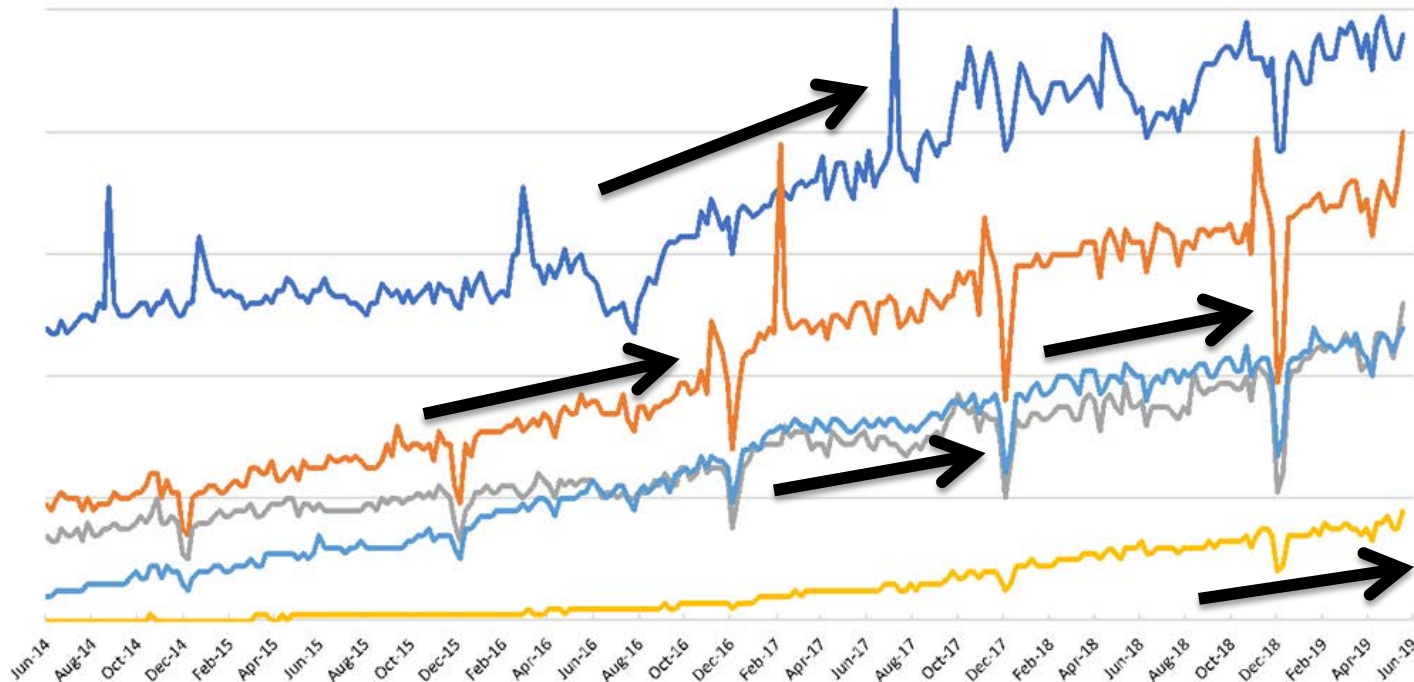→ **Big Data**
→ **Cloud Computing (Search Term)**

Some smaller areas: Google Trends last 5 years *(Topics unless stated)*

➔ **Cloud Computing (Search Term)**

➔ Grid Computing

➔ **HPC**

➔ **Edge Computing**

➔ **SuperComputing (Search Term)**

Some growing areas: Google Trends last 5 years *(Topics unless otherwise stated)*

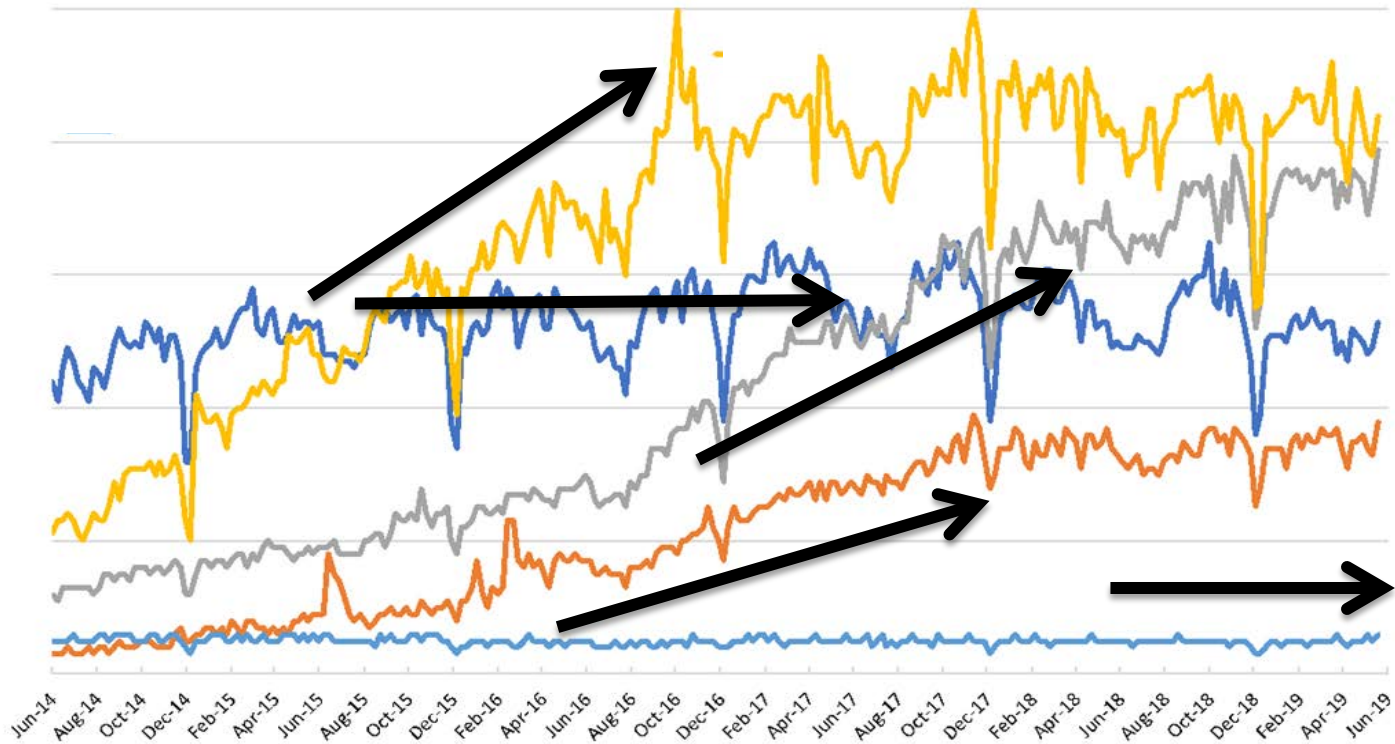→ **Artificial Intelligence**

→ **Amazon Web Services**

→ **Azure (Comp. App)**

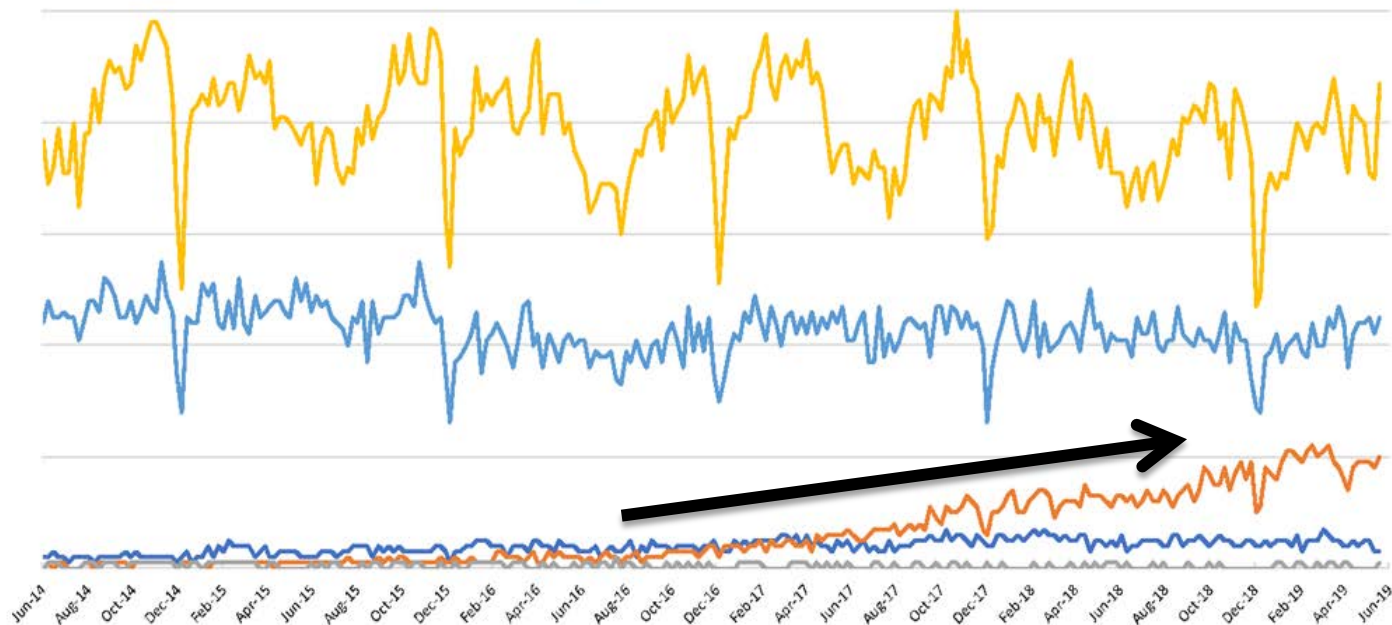→ **Docker (Software)**

→ **Kubernetes (Comp. App)**

Some medium size areas: Google Trends last 5 years *(Topics unless otherwise stated)*

AI IS 10X BIG DATA, CYBERINFRASTRUCTURE, EXASCALE SMALL

→ **Internet of Things**

→ **Machine Learning (Search Term)**

→ **Big Data**

→ **Deep Learning**

→ **High Performance Computing (HPC)**

Some smaller areas: Google Trends last 5 years *(Topics unless otherwise stated)*

→ **Parallel Computing (Programming Paradigm)**

→ **HPC**

→ **Edge Computing**

→ **Fog Computing**

→ **Cyberinfrastructure**