The Potential of Unstructured and Big Data for Public Health

Ian Brooks, PhD
Center for Health Informatics
School of Information Sciences
University of Illinois
ianb@Illinois.edu

4-6 December Atlanta, Georgia, USA





What would you do differently if computing were unlimited?



NCSA Blue Waters



13.34 Pflop

(13,340,000,000,000,000 floating point operations / second)

396,040 AMD processor cores 4,224 NVIDIA GPUs

1.476 PB System Memory

26.4 Pbytes Primary Storage > 1 TB/s aggregate I/O Bandwidth

Up to 0.5EB Near Line Storage 58 GB/s



NCSA Blue Waters in Context



13.34 Pflop

Multiplying two 15-digit numbers together every second would take more than 422 million years

1.476 PB System Memory

- ~190,000 DIMMS
- ~ 300 million dSLR images

26.4 Pbytes Primary Storage

> 17,000 Disks

Enough to store all the printed documents in all of the world's libraries

Up to 0.5 EB Near Line Storage

Enough to store 10% of all the words spoken in the existence of humankind



TUKALEIDOSCOPE ATLANTA 2019



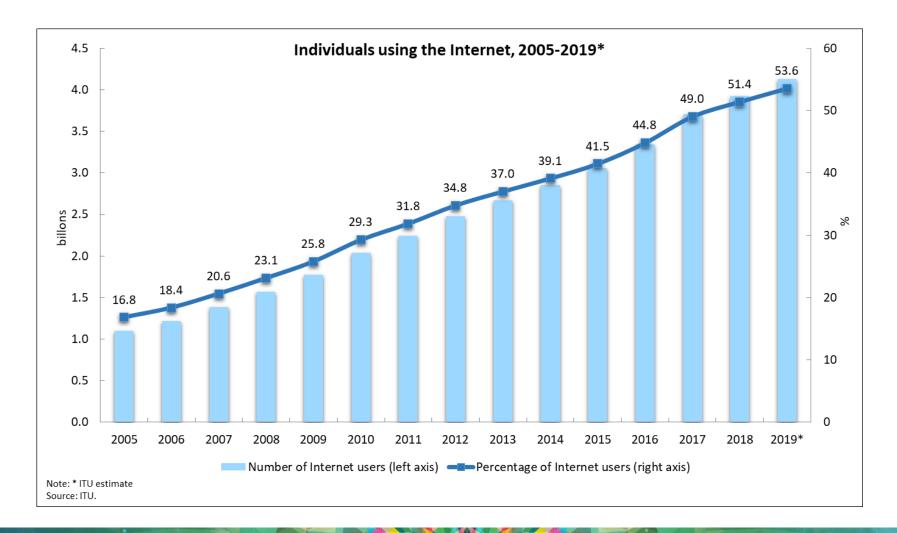




For all practical purposes in public health computing IS unlimited









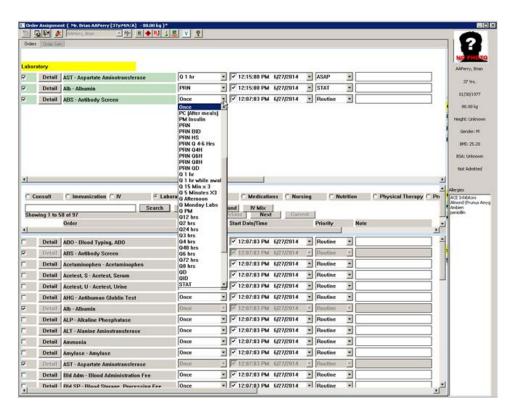


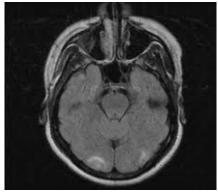


~2.8 Billion Users

www.internetworldstats.com

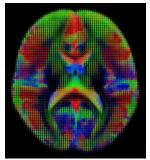






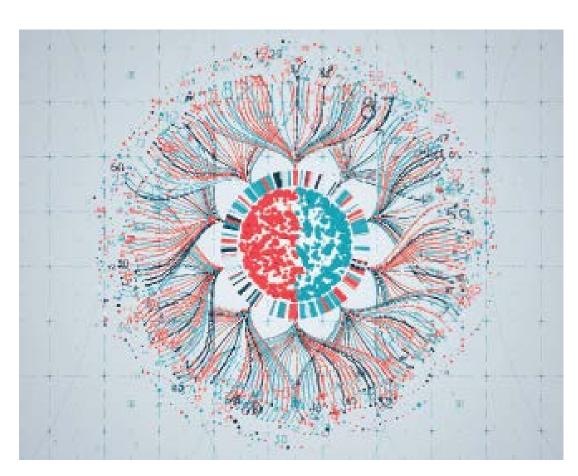








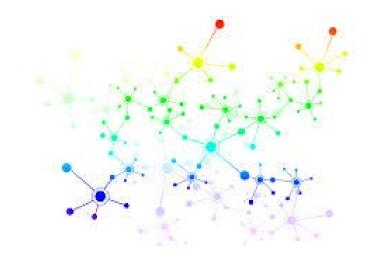










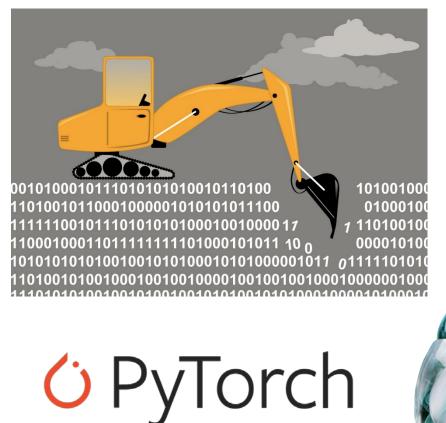


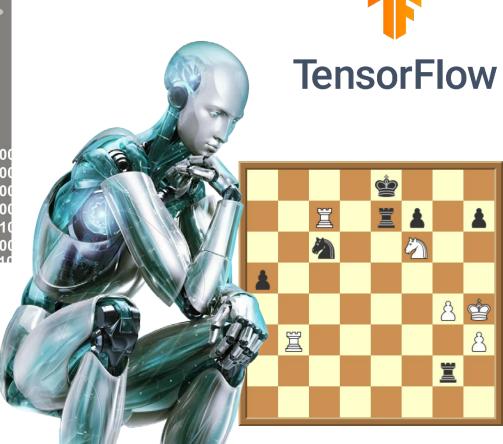




















How do you measure the health of a population?



"There's something going around"





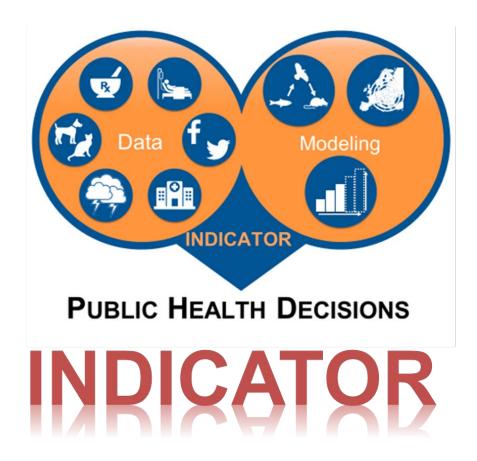
TUKALEIDOSCOPE ATLANTA 2019













TUKALEIDOSCOPE ATLANTA 2019









TUKALEIDOSCOPE ATLANTA 2019

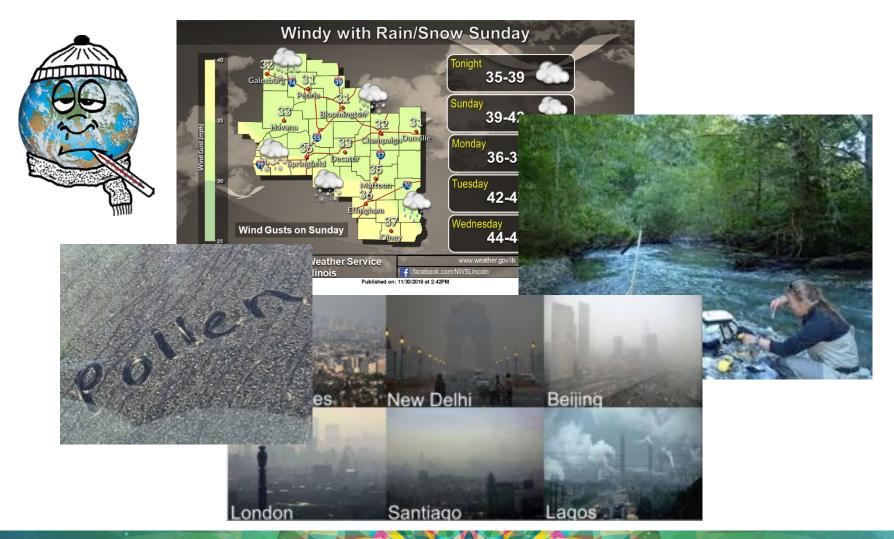




TUKALEIDOSCOPE ATLANTA 2019



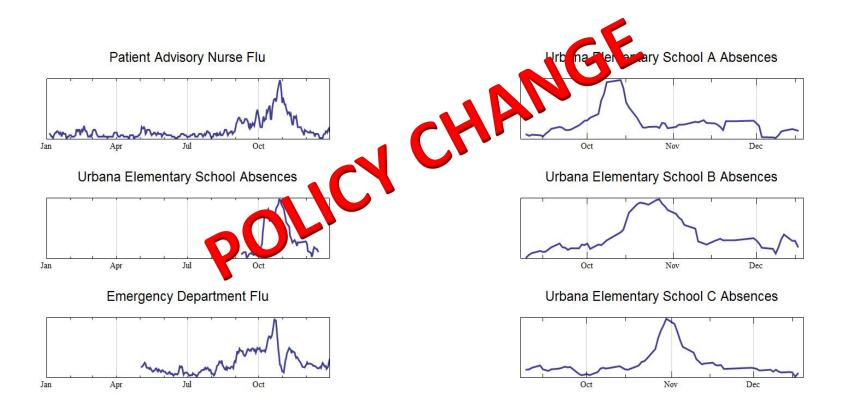






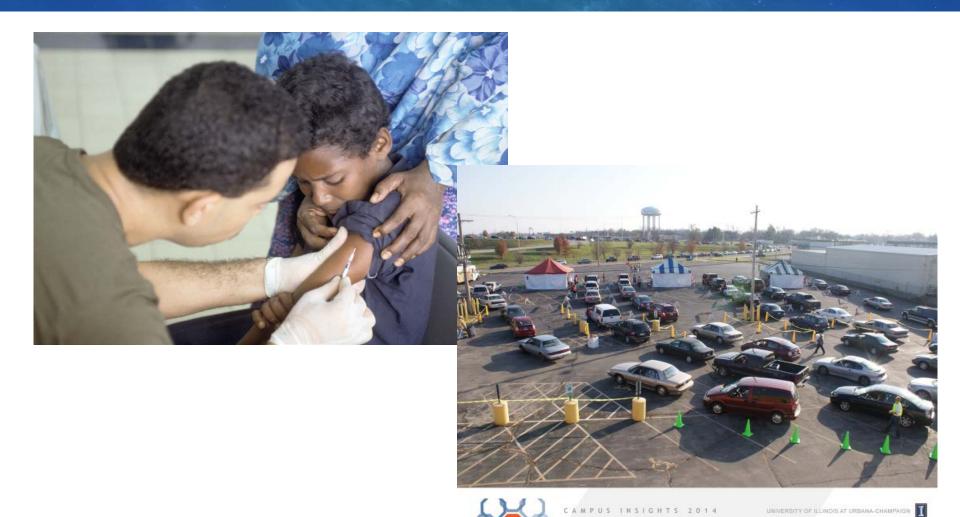


H1N1 Outbreak 2009





TUKALEIDOSCOPE ATLANTA 2019





TUKALEIDOSCOPE ATLANTA 2019

Lessons learned

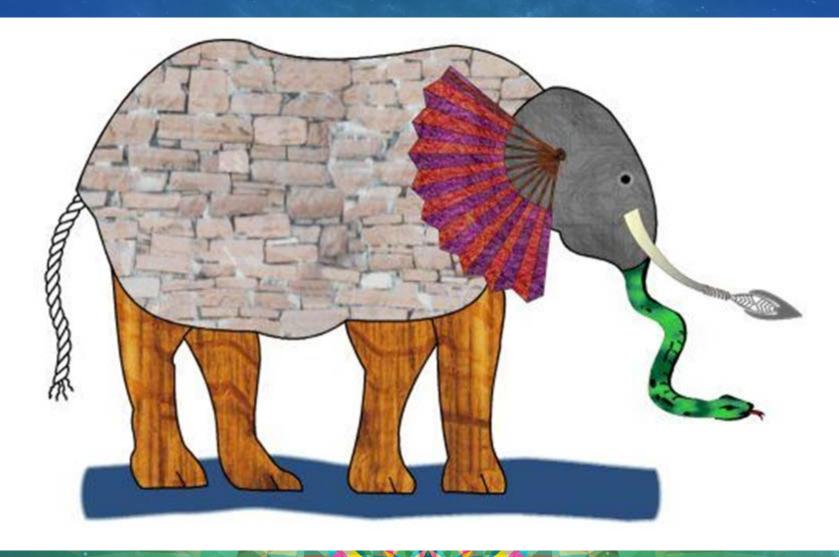








TUKALEIDOSCOPE ATLANTA 2019







Final Thoughts

Technical Engineering of Health Informatics projects is easy

At least when compared to the required Social Engineering



Thank you