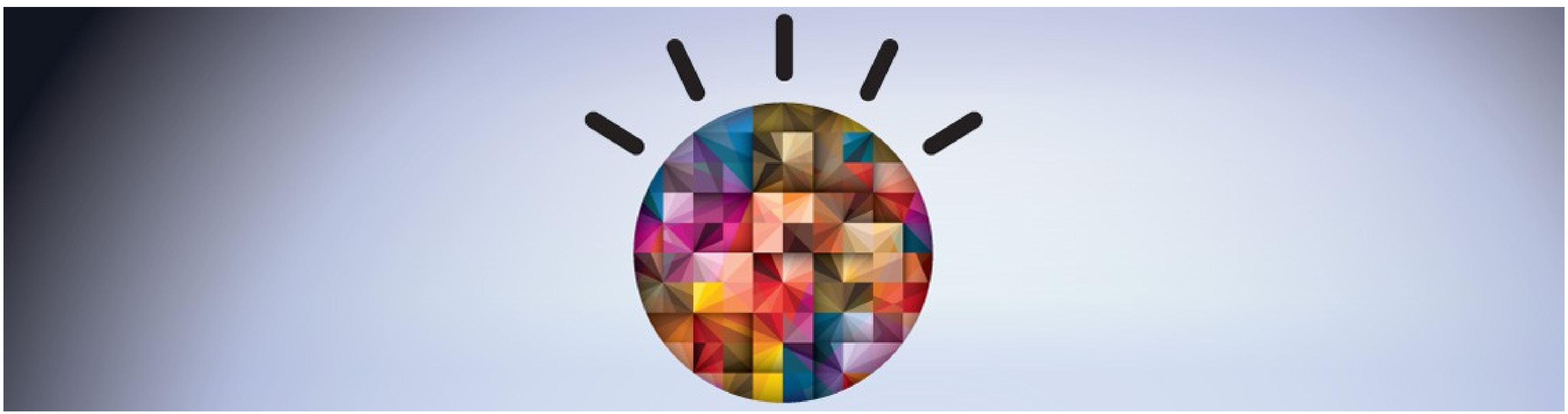
17 July 2013
IEEE NAS Conference
Xi'An China

The Big Deal about Big Data – a perspective from IBM Research

H. Peter Hofstee (& Kevin J. Nowka)
IBM Research -- Austin



Addressing BIG Problems with BIG Data

Sources and Types of Data

Classifying and Processing Data

A bit about Systems

In 2011 US healthcare spending was
$2.7 Trillion -- 17.9% of GDP.
IOM estimates $750B wasted due to unnecessary and inefficiently delivered services, excessive administrative costs, missed prevention opportunities, and fraud.

Src: Center for Medicare & Medicaid Services, Institute of Medicine

In 2011 US healthcare spending was
$2.7 Trillion -- 17.9% of GDP.
IOM estimates $750B wasted due to unnecessary and inefficiently delivered services, excessive administrative costs, missed prevention opportunities, and fraud.

In 2012 CMS stored 400 Terabytes of Medicare/Medicaid claims data – and estimated this to double by YE2015.
US Healthcare data is estimated to exceed 150 Exabytes (150 Million TB).

Src: Center for Medicare & Medicaid Services, Institute of Medicine; Trenkle, Tackling Future CMS Requirements; , Roger Foster, “How to Harness Big Data for Improving Public Health,” Government Health IT, April 3, 2012, at http://www.govhealthit.com/news/how-harness-big-data-improving-public-health

Healthcare industry is beset with some of the most complex information challenges we collectively face

- Medical information is doubling every 5 years, much of which is unstructured
- 81% of physicians report spending 5 hours or less per month reading medical journals

1 in 5

diagnosis that are estimated to be inaccurate or incomplete

1.5 million

errors in the way medications are prescribed, delivered and taken in the U.S. every year

44,000 -98,000

# of Americans who die each year from preventable medical errors in hospitals alone

"Medicine has become too complex. Only about 20% of the knowledge clinicians use today is evidence-base." Leading Chief Medical & Scientific Officer

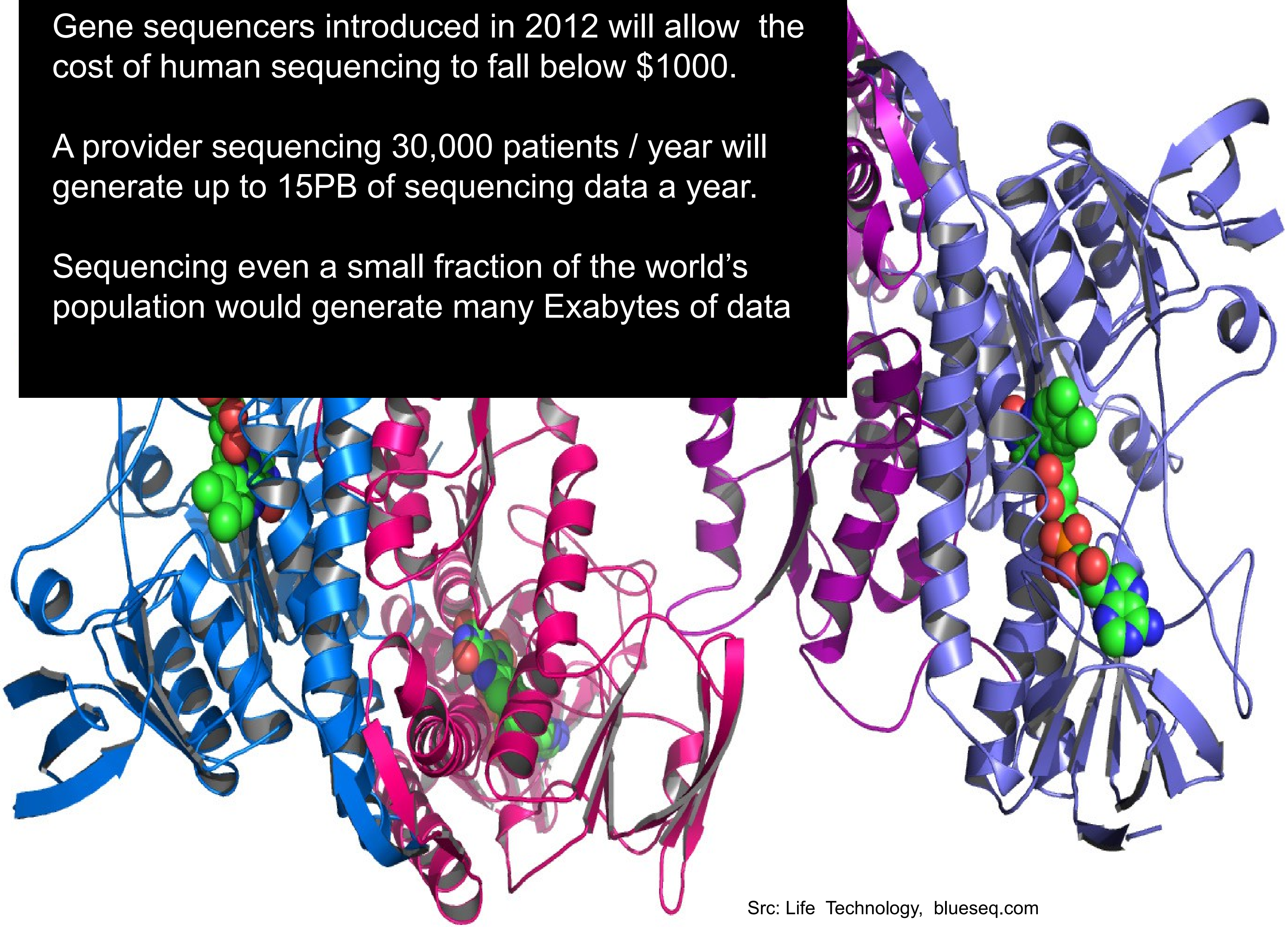
Big Data Analytics in Smarter Hospitals

Big Data enabled doctors from University of Ontario to apply neonatal infant monitoring to predict infection in ICU 24 hours in advance

Gene sequencers introduced in 2012 will allow the cost of human sequencing to fall below $1000.

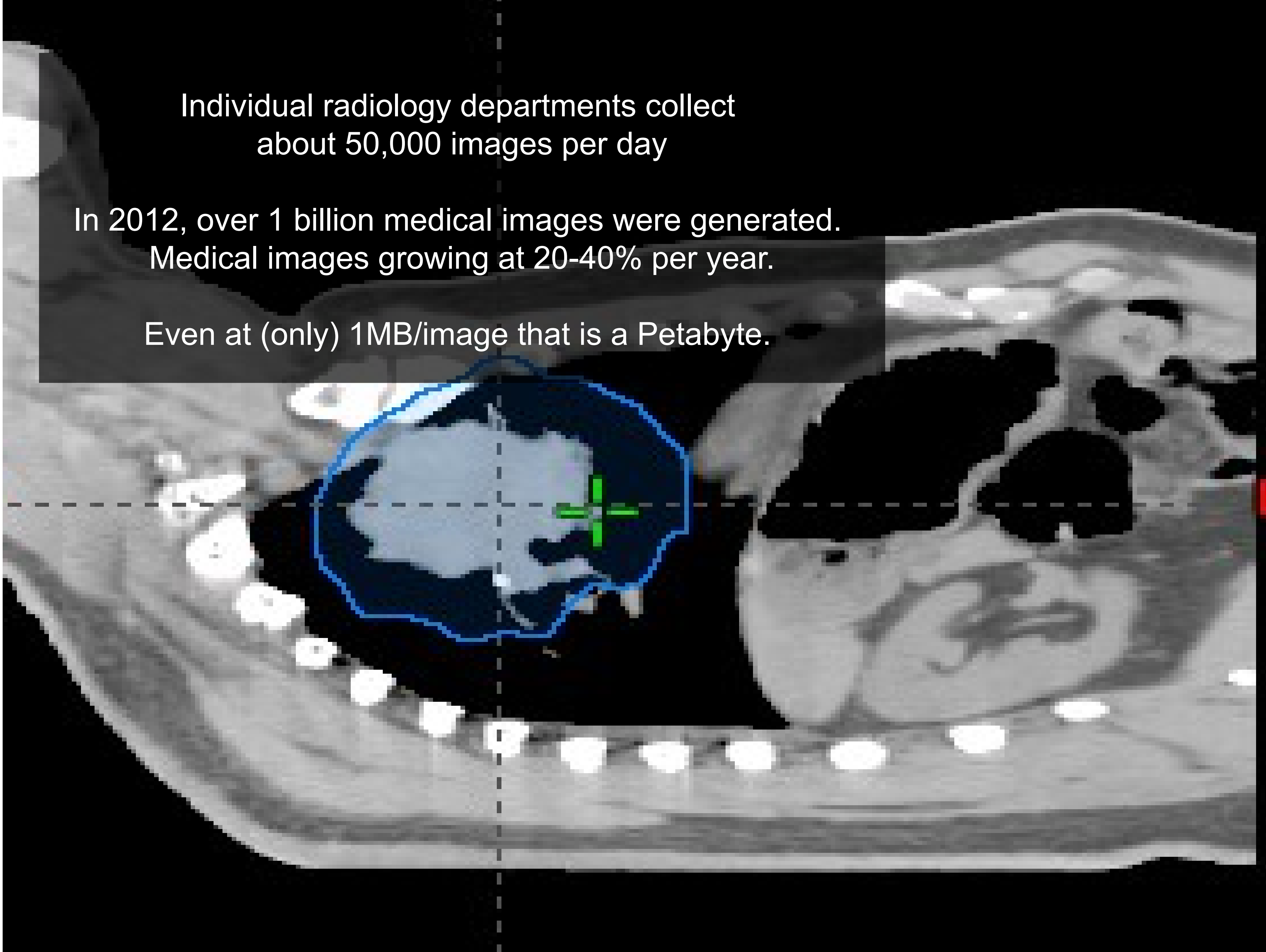
A provider sequencing 30,000 patients / year will generate up to 15PB of sequencing data a year.

Sequencing even a small fraction of the world's population would generate many Exabytes of data



Src: Life Technology, blueseq.com

Image: Midwest Center for Structural Genomics

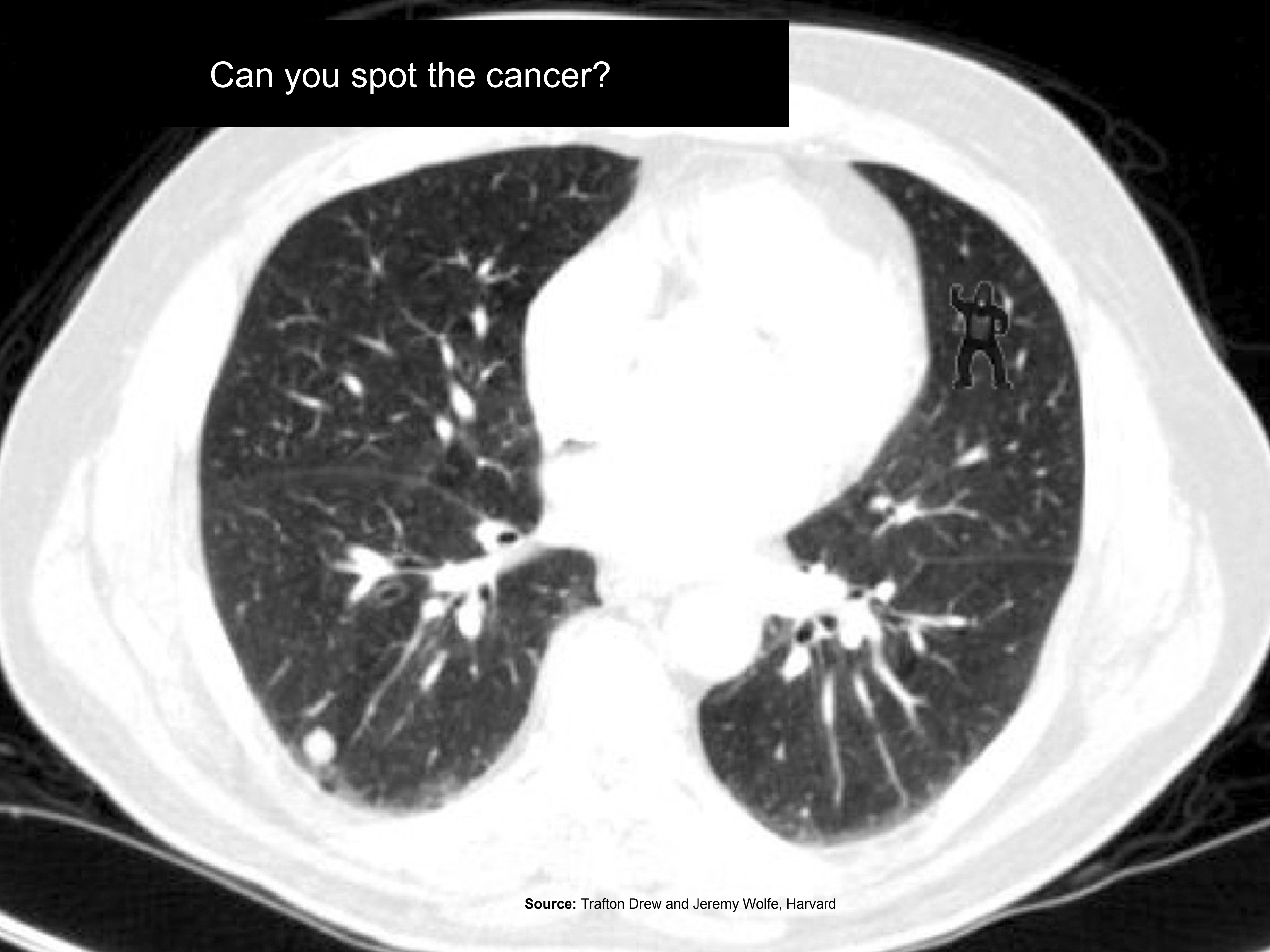


Individual radiology departments collect about 50,000 images per day

In 2012, over 1 billion medical images were generated. Medical images growing at 20-40% per year.

Even at (only) 1MB/image that is a Petabyte.

Can you spot the cancer?



Source: Trafton Drew and Jeremy Wolfe, Harvard

“Inattentional blindness“ led 83% of radiologists searching this image for cancerous nodules to miss the gorilla.

Source: Trafton Drew and Jeremy Wolfe, Harvard

Daily Retina Scan for Everyone

6B x 1MB x 365/y = ~2 Trillion MB = ~2 Exabytes/year

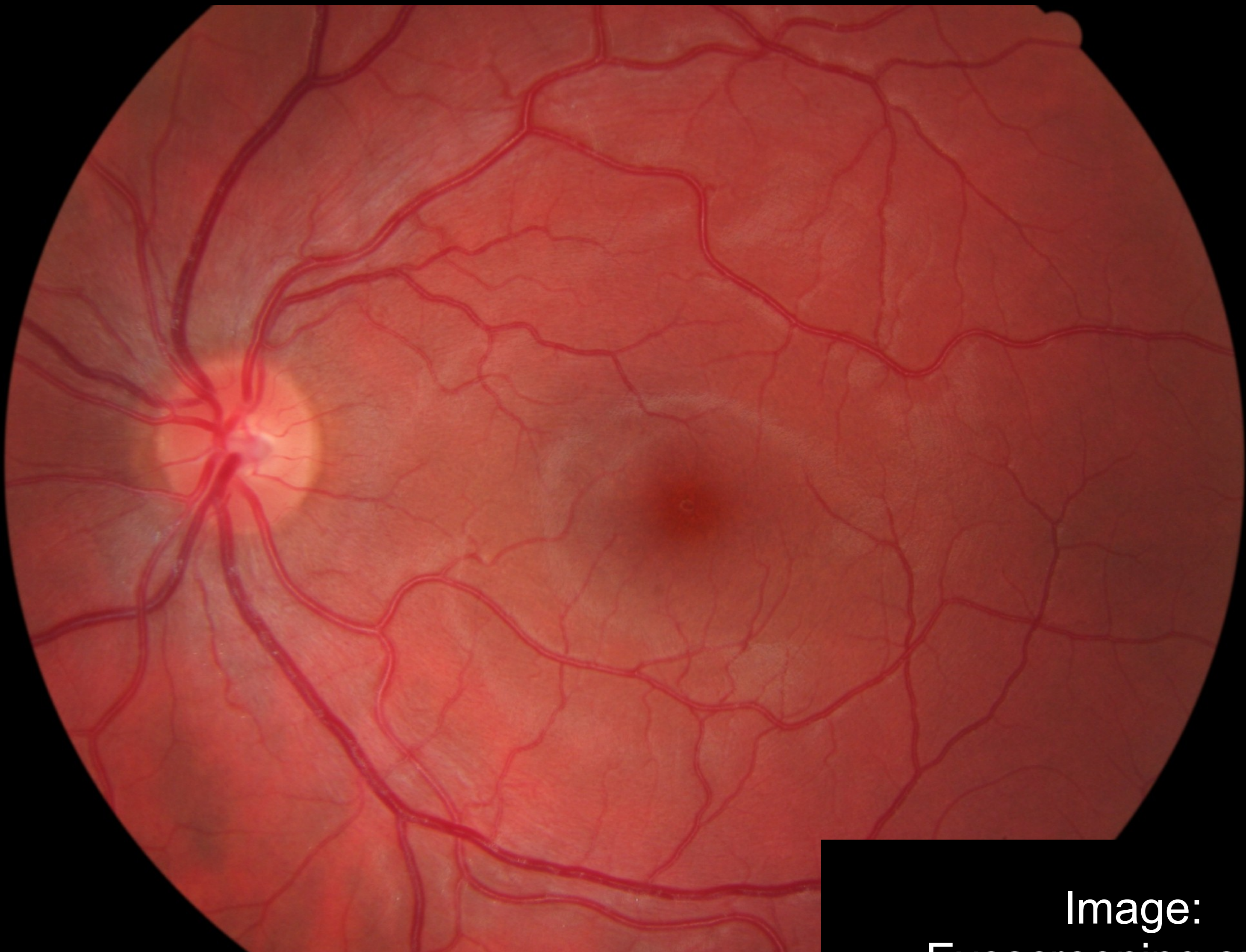
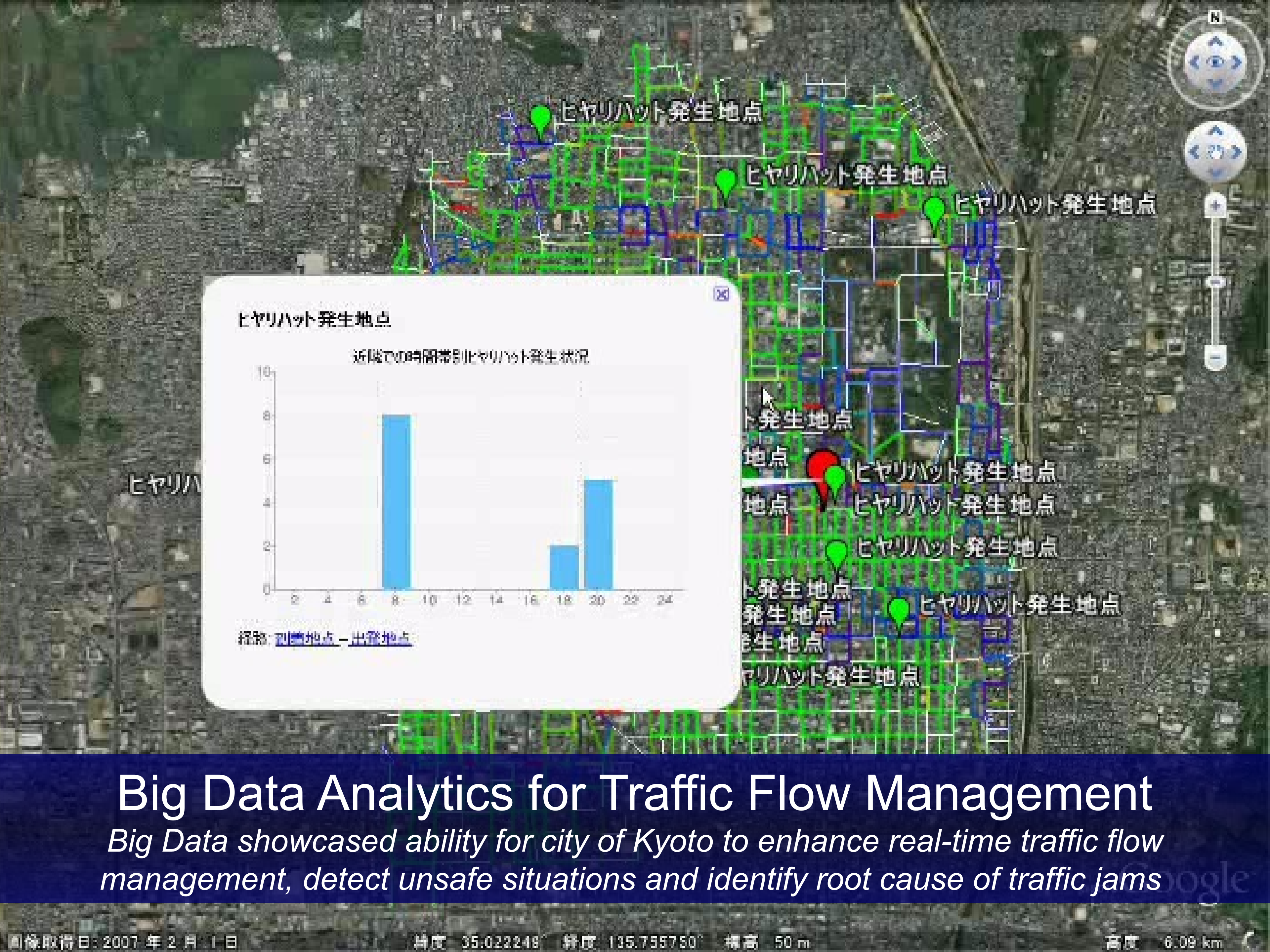


Image:
Eyescreening.org.uk


Congested urban roadways cost US $121 billion annually in the form of 5.5 billion lost hours and
2.9 billion gallons of wasted gas….
with 56 Billion pounds of additional
emitted carbon dioxide.
CAR WASH
Power
Src: TTI ’13 Urban Mobility Report



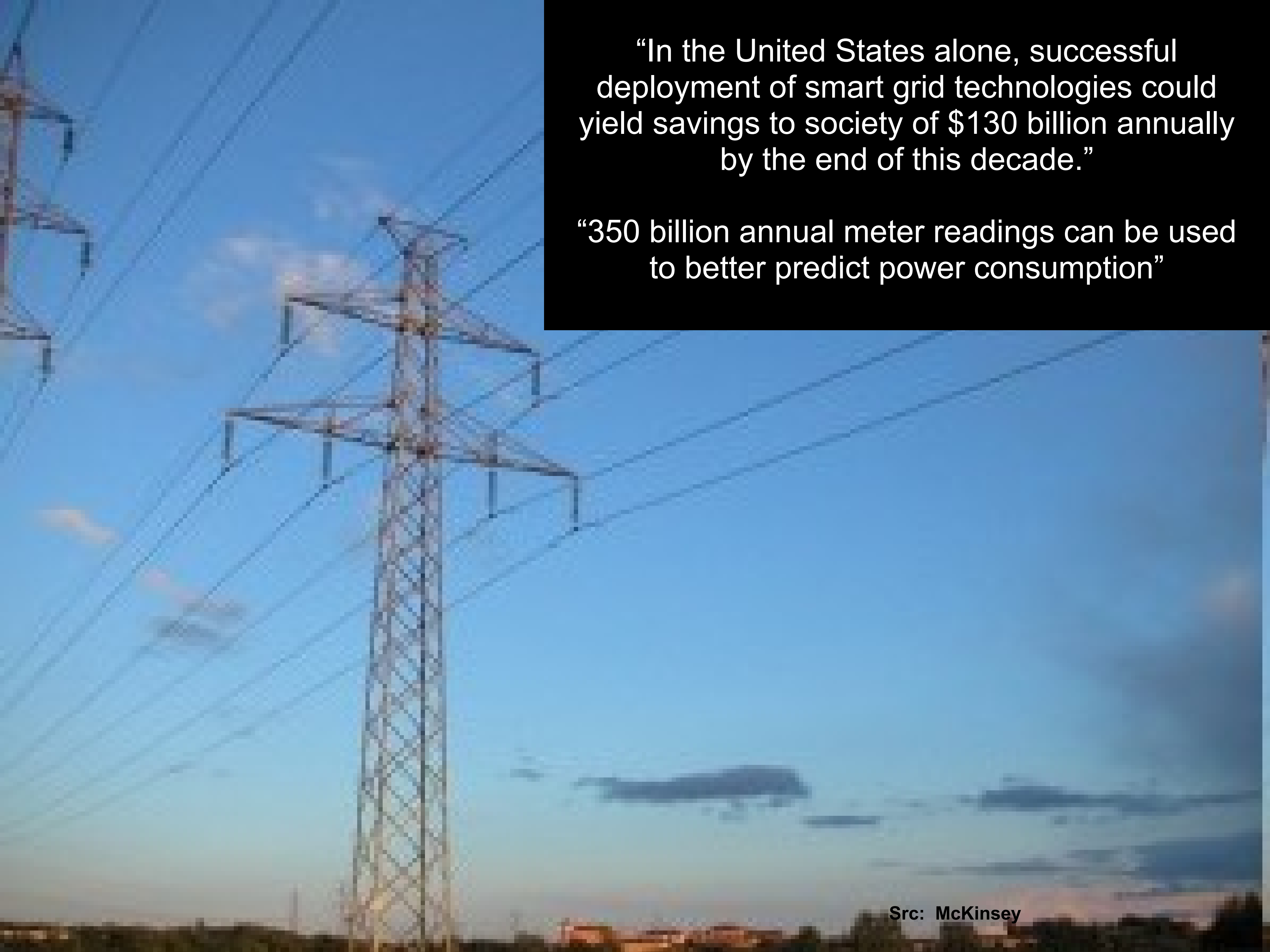
Big Data Analytics for Traffic Flow Management

Big Data showcased ability for city of Kyoto to enhance real-time traffic flow management, detect unsafe situations and identify root cause of traffic jams

An increasingly sensor-enabled and instrumented environment generates HUGE volumes of data with MACHINE SPEED characteristics…

1 BILLION lines of code

EACH engine generating 10 TB every 30 minutes!



"In the United States alone, successful deployment of smart grid technologies could yield savings to society of $130 billion annually by the end of this decade."

"350 billion annual meter readings can be used to better predict power consumption"

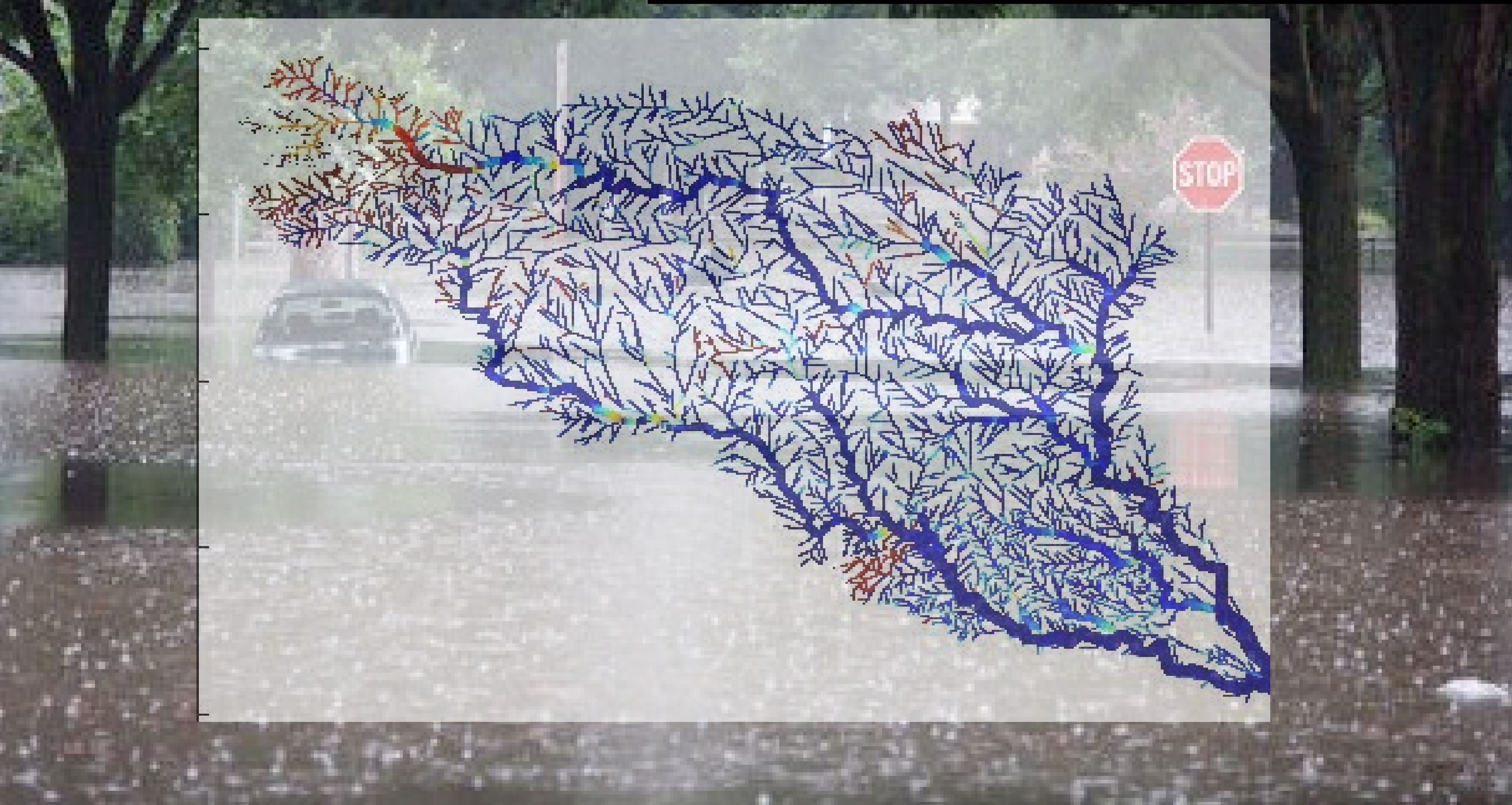
Src: McKinsey

Central Texas is among the world's most flash flood prone locations. In 24 hours 8-9 Sept 1921, over 38 inches of rain fell over Thrall, TX. 215 people died in that storm. Despite >500 stream flow gages in "Flash Flood Alley," and access to millions of historical readings there have been more than 200 flood-related fatalities since 1996

Src: Lott, Monthly Weather Review 1955; Maidment 2010 TFFC Workshop

UT Austin / IBM Austin Watershed Simulation

Ultimately need to bring real-time measured data, historical data, and modeled data together.



Vestas models weather to optimize placement of turbines, maximizing power generation and longetivity based on 2.5 Petabytes of information.

- Public wind data is available on 284km x 284 km grids (2.5o LAT/LONG)
- Perspective: The Vestas Wind library, as HD TV would take 70 years to watch
- More data means more accurate and richer models (adding hundreds of variables)
  - Granularity 27km x 27km grids: driving to 9x9, 3x3 to 10m x 10m simulations
  - Reduce time required to identify placement of turbine from weeks to hours.

Big Applications for BIG Data Analytics

Neonatal Care

Trading Advantage

Environment

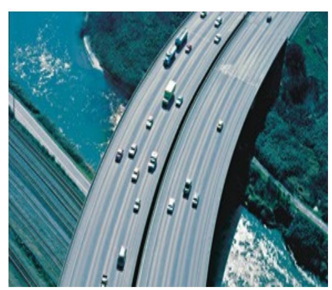
Law Enforcement

Radio Astronomy

Telecom

Manufacturing

Traffic Control



Fraud Prevention

Merging the Traditional and Big Data Approaches

Traditional Approach

Structured & Repeatable Analysis

Business Users

Determine what question to ask

IT

Structures the data to answer that question

- Monthly sales reports
- Profitability analysis
- Customer surveys

Big Data Approach

Iterative & Exploratory Analysis

IT

Delivers a platform to enable creative discovery

Business

Explores what questions could be asked

- Brand sentiment
- Product strategy
- Maximum asset utilization

Addressing BIG Problems with BIG Data

Sources and Types of Data

Classifying and Processing Data

A bit about Systems

“Big Data” has come to mean drawing value from data which has the following characteristics:

Volume:	Scale from Petabytes (1000 Terabytes) to Exabytes (million Terabyte) to Zettabytes (billion Terabyte)
Variety:	Complex data in many different formats from many sources
Velocity:	Streaming data requiring fast response
Veracity:	Trust improves as the number/variety of data grows

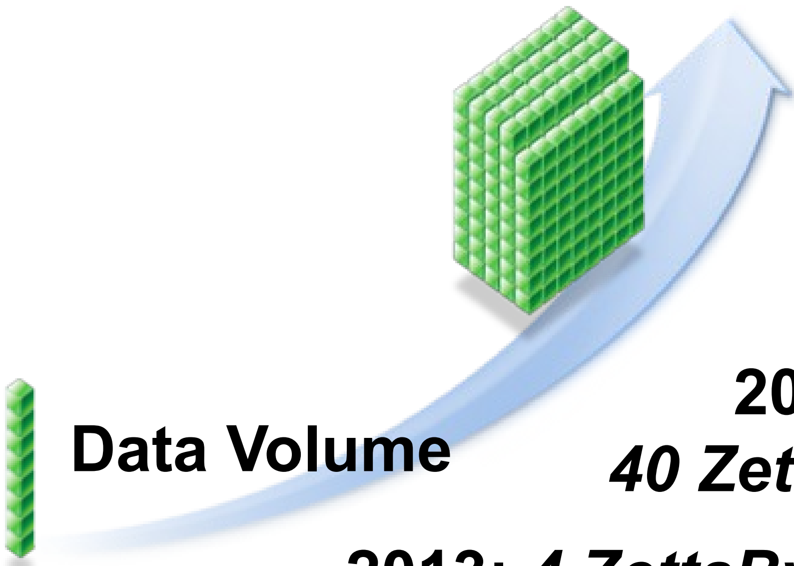
Big Data Volume: Digital Data is expected to double every 2 years

50x

as much Data and Content
Over Current Decade

LIBRARY OF CONGRESS

2.5 Exabytes (million terabytes)
per day being created …
800-times the estimated size of the collection in the
US Library of Congress
(2 LoC estimate from 1997)



Data Volume

2020:
40 ZettaBytes

2013: ***4 ZettaBytes***

2009: 0.8 ZettaBytes (***Billion Terabytes)***

Source: The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East,; IDC Predictions 2013
http://event.lvl3.on24.com/event/54/34/13/rt/1/documents/slidepdf/wc20130108.pdf

According to the Sloan Digital Sky Survey,
“the currently observable universe is home to of the order of 6 x 10^22 [60 billion trillion] stars“

By 2017 the total bits of digital data will surpass the number of stars in the observable universe!

Data increasing as access increases:
In 2012 an estimated 2.4 billion people had Web access ...
…including 530 million Web users in China.
Two-thirds of the world's population does not have access

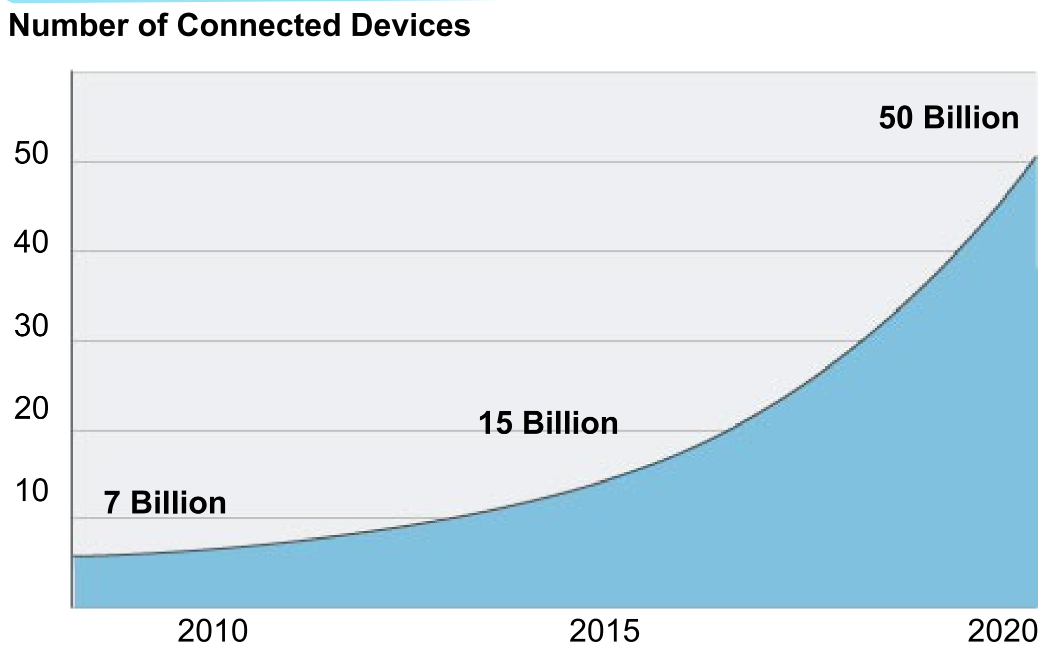
Worldwide mobile telephone subscriptions reached 6.3 billion in 2012.
5-times as much Data as Voice traffic
5.9 trillion text messages were sent in 2011.

Src: Ericsson 2013 Interim Mobility Report; Informa

Big Data Variety

An increased variety of data sources are generating large quantities of data …

Connected devices are driving much of the data growth: Security sensors, cameras, light bulbs, refrigerators, utility sensors, health sensors, tablets, netbooks, eBook readers, Internet TVs, digital picture frames, cars....

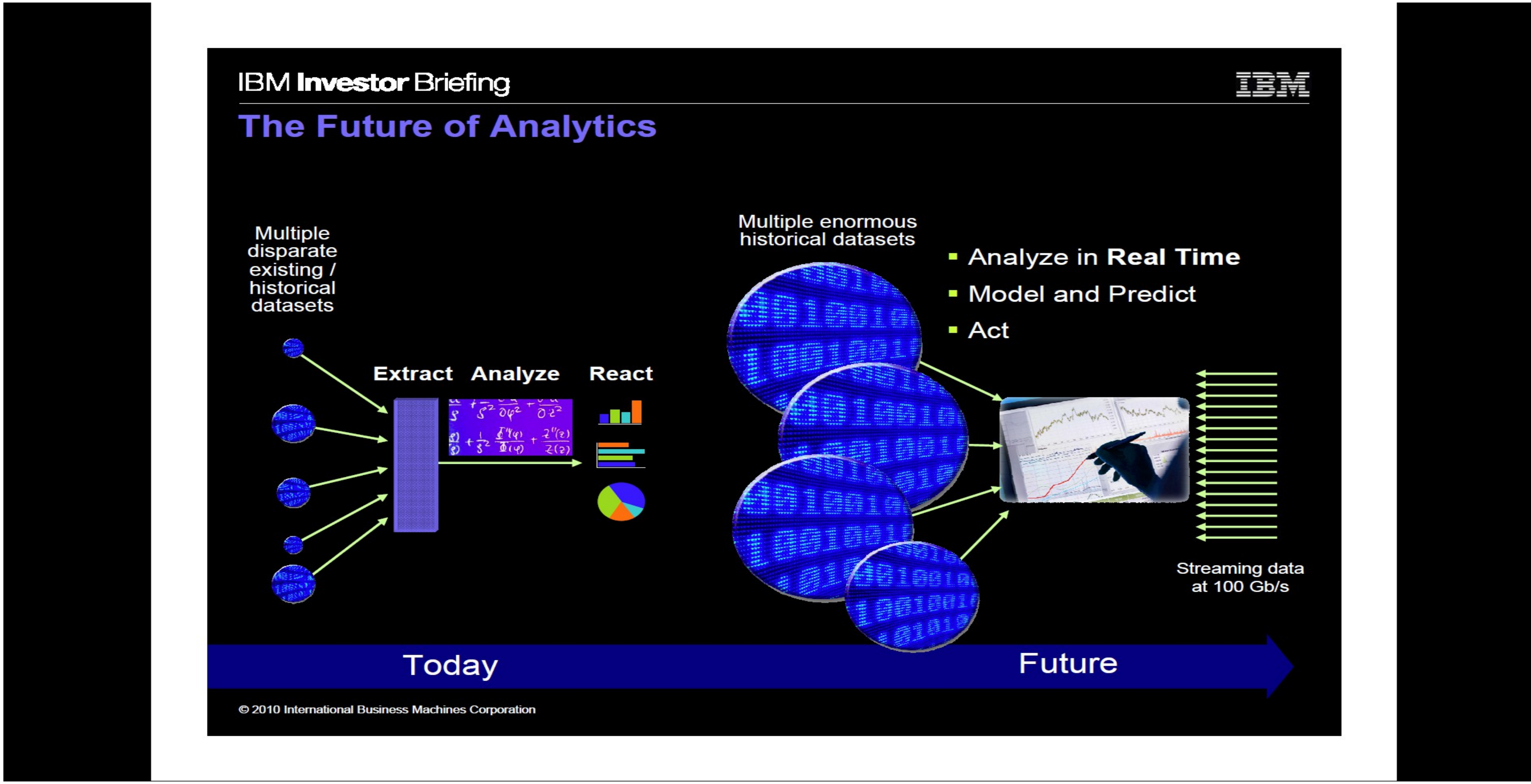


Multiple Sources: Intel, Ericsson, Gartner, etc.

Connected device annual growth rate in Security, Health care, and Utility sectors exceed 45%.

CAGR for all connected devices is 35% 2010-15

Big Data Velocity



Data at Rest

Data in Motion

Big Data Velocity: Analytics on Data in Motion

Telco Promotions
6B records/day
10 ms/decision
270TB for Deep Analytics

Smart Traffic
250K GPS probes/sec
630K segments/sec
2 ms/decision, 4K vehicles

SRC: IBM 2012 GTO

13.345567% of statistics used in Big Data talks are complete fabrications ☺

How do you know I’m telling the truth?

Big Data Veracity

1 in 3

Business leaders frequently make decisions based on information they don't trust, or don't have

1 in 2

Business leaders say they don't have access to the information they need to do their jobs

Data that is incomplete, inconsistent, missing, incorrect, ambiguous, too late, fake, corrupted…can be disastrous

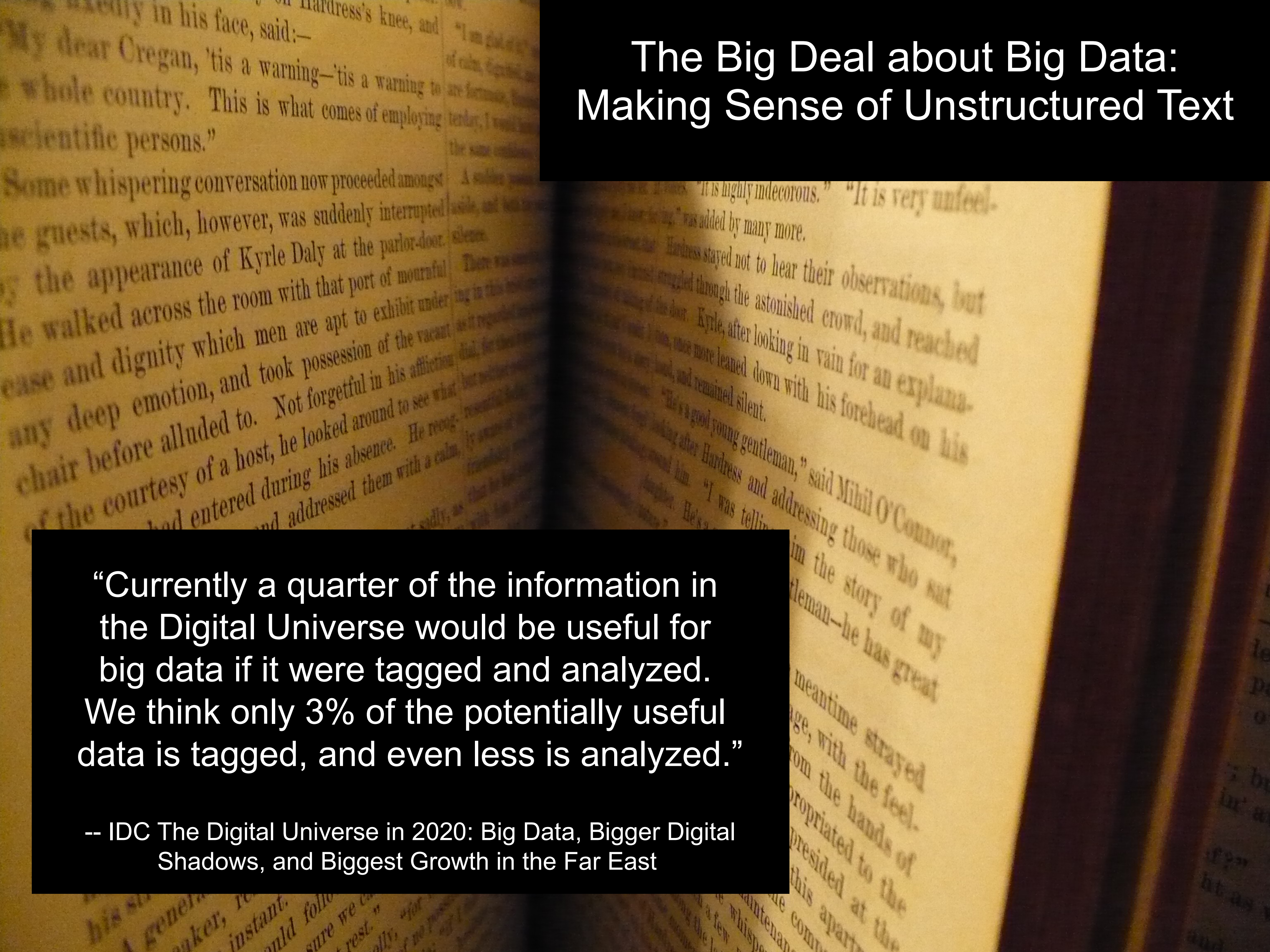
SRC: IBM Institute for Business Value http://www.ibm.com/services/us/gbs/thoughtleadership/

Addressing BIG Problems with BIG Data

Sources and Types of Data

Classifying and Processing Data

A bit about Systems



The Big Deal about Big Data: Making Sense of Unstructured Text

"Currently a quarter of the information in the Digital Universe would be useful for big data if it were tagged and analyzed. We think only 3% of the potentially useful data is tagged, and even less is analyzed."

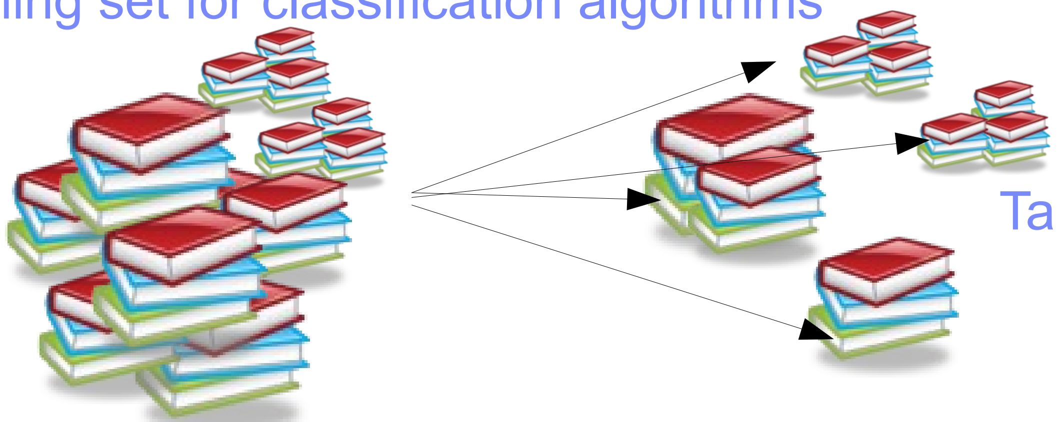
-- IDC The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East

Leveraging the Wisdom of the Crowd ...

Trusted sources of classification

LIBRARY OF CONGRESS

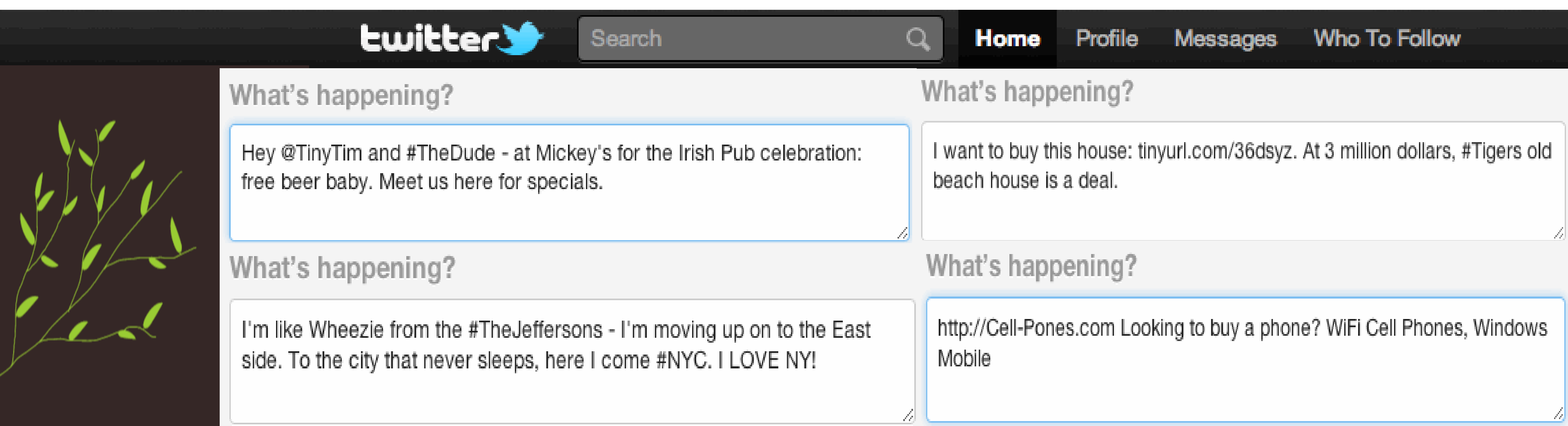
Training set for classification algorithms



Tagged documents

e.g. Gattiker e.a., IBM Journal 57 3/4

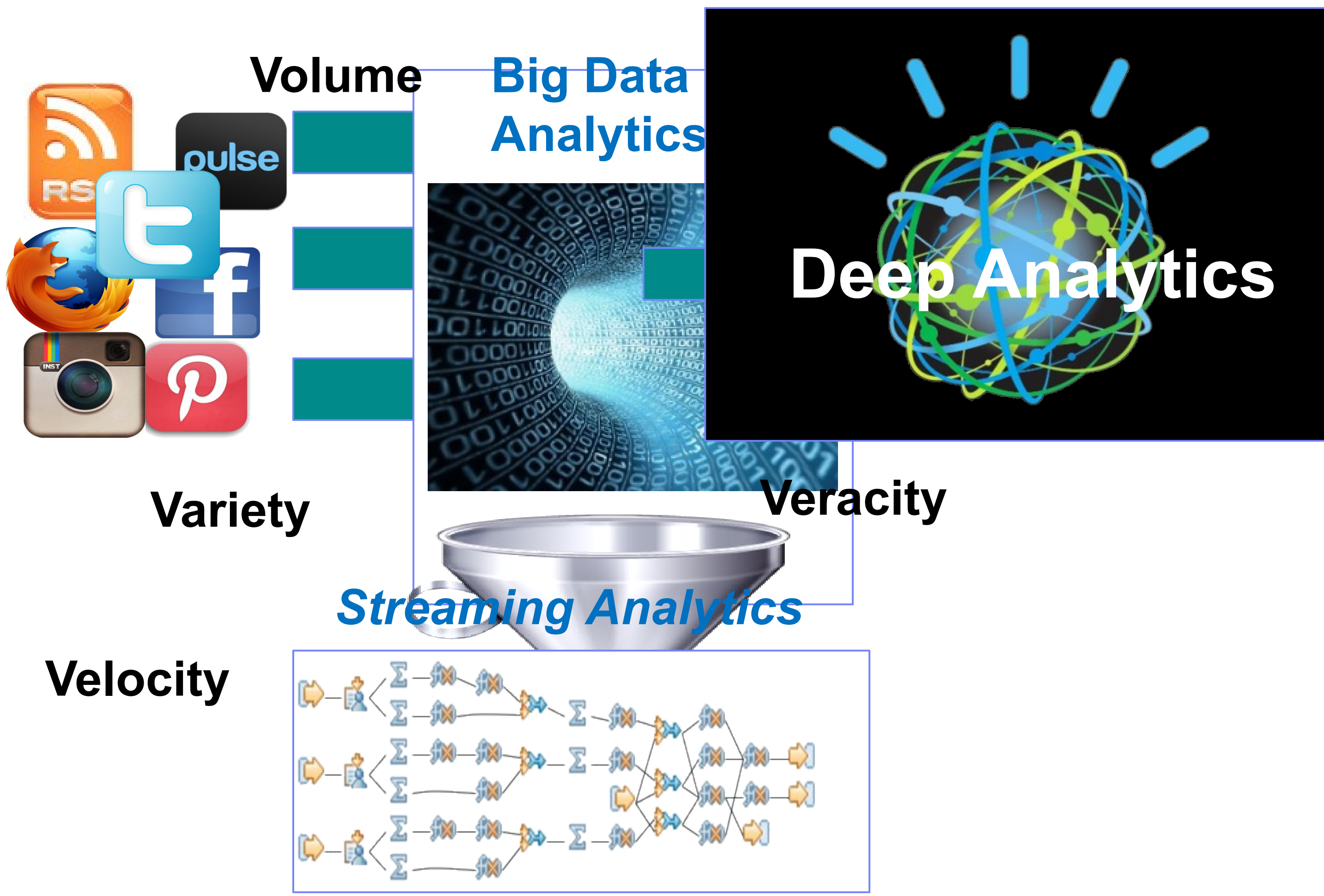
Unstructured text analytics to extract intent, drive sentiment analysis



Unstructured text analytics to extract intent, drive sentiment analysis

Natural Language Access

IBM WATSON

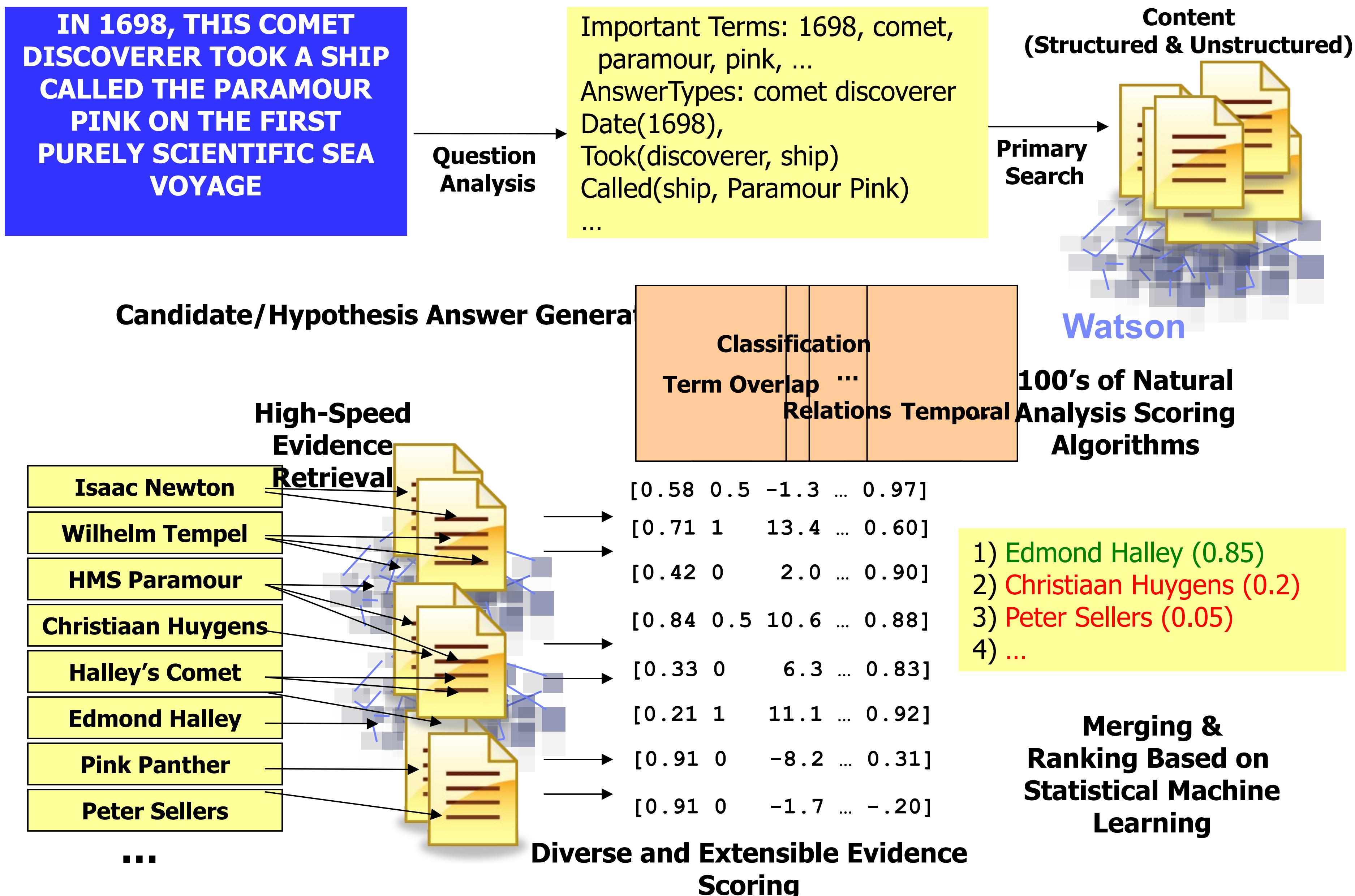



THINK
सोचिए
ДУМАЙ
$1,000
KEN
$2,000
WATSON
$1,200
BRAD

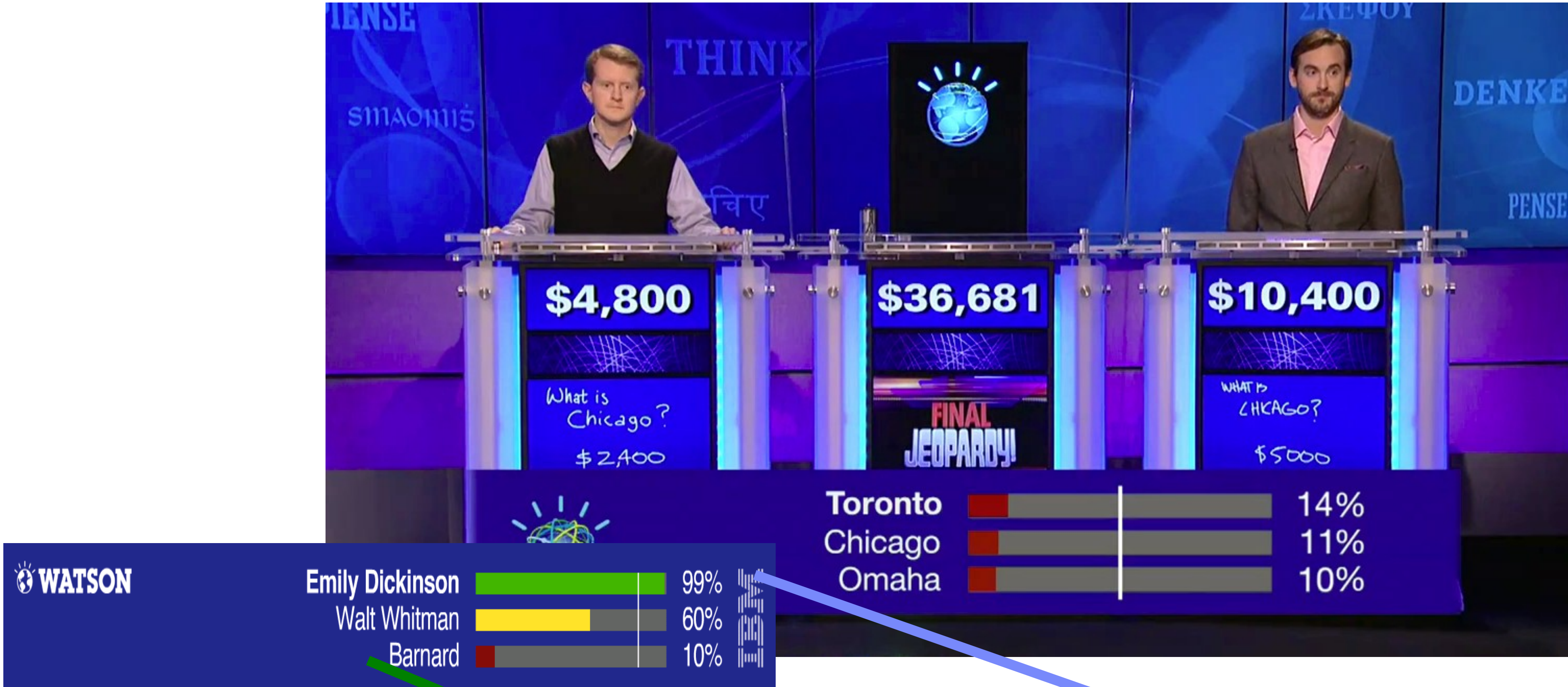
What is Watson?
Automatic Open-Domain Question Answering System

- Given
  - Rich **Natural Language Questions**
  - Over a **Broad Domain of Knowledge**
- Deliver
  - **Precise Answers:** Determine what is being asked and provide precise responses
  - **Accurate Confidences:** Determine *likelihood answer is correct*
  - **Consumable Justifications:** Explain why the answer is right
  - **Fast Response Time:** Precision & Confidence in <3 seconds

Watson answers by finding, reading, scoring and combining evidence



There is power in knowing when you don’t know the answer!

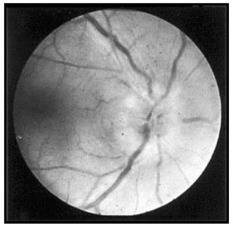


“There are **known knowns**; there are things we know we know.
We also know there are **known unknowns**; that is to say,
we know there are some things we do not know.
But there are also **unknown unknowns** – the ones we don’t know we don’t know.”
-- Donald Rumsfeld, US Secretary of Defense

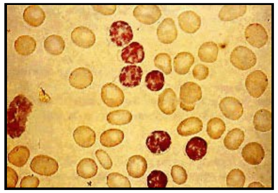
Medical Image Category Recognition

- **With billions of medical images, use the computer to categorize them.**
  - Recognizing up to millions of categories of images in medicine
  - Massive data-driven modeling – extract large number of visual features and learn discriminative models from massive training data

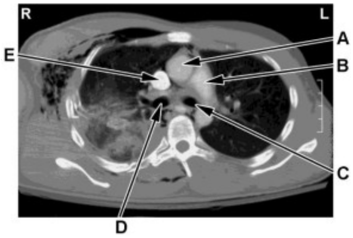
- **Massive data:**
  - Training examples (need 100's per category) (= ~10M – 100M images)
  - Sources = textbooks, journals, repositories, open data sets, Web



- **Classification schemes:**
  - Construct large taxonomy (across modality, anatomy, view, pathology) (~1M)
  - Inform from known medical ontologies, textbooks, references



- **Visual classifiers:**
  - Train discriminative ensemble visual feature classifiers
  - Learn multi-modal classification where applicable
  - (e.g., visual features + text captions, annotations, related text, patient record

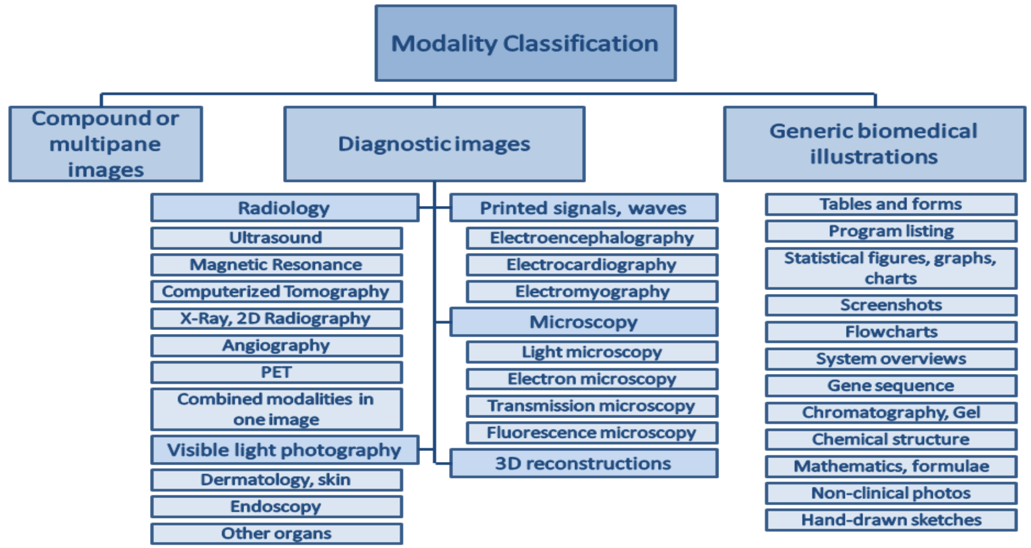


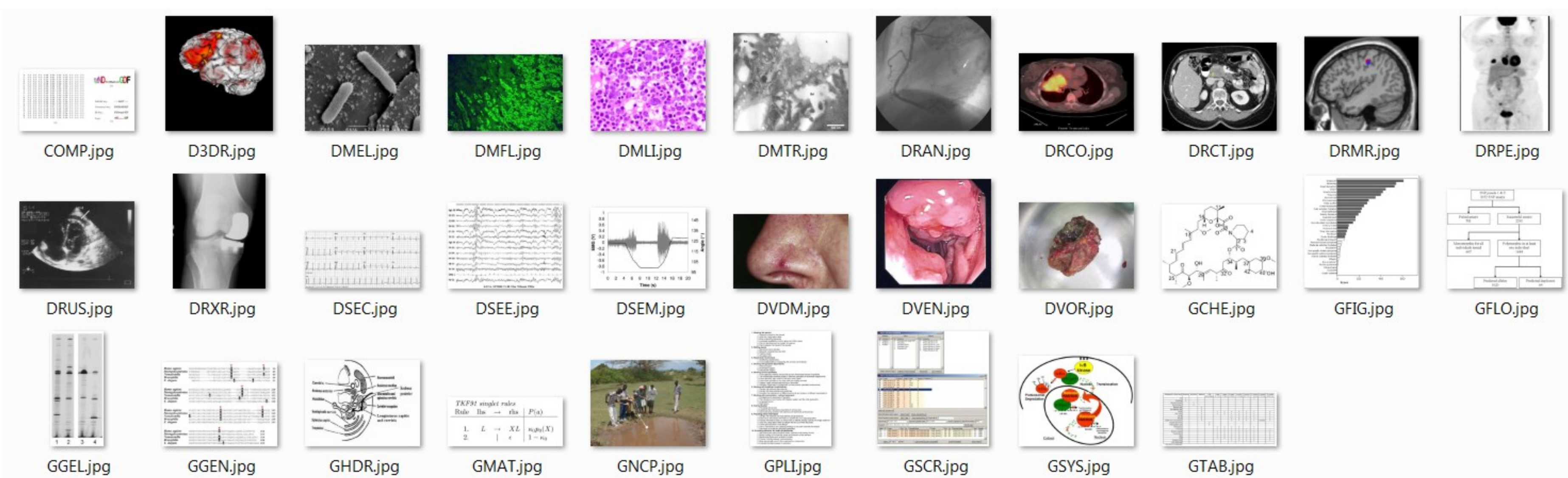
Medical Imaging Classification

IBM

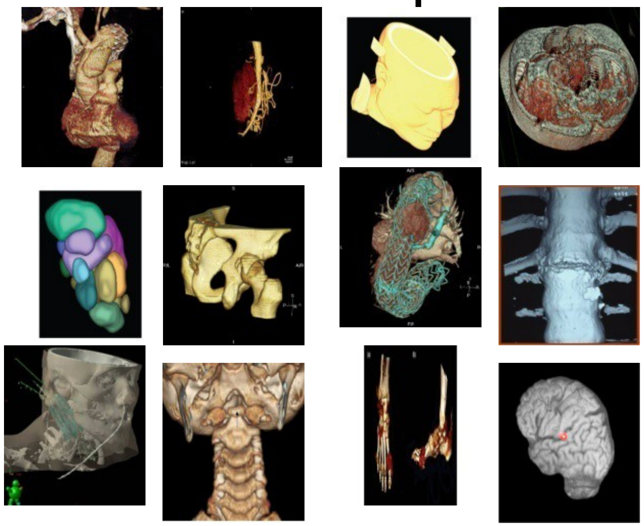
ImageCLEF 2012

After training on 1026 images, computer attempts to classify 1200 Test Images into 31 Categories

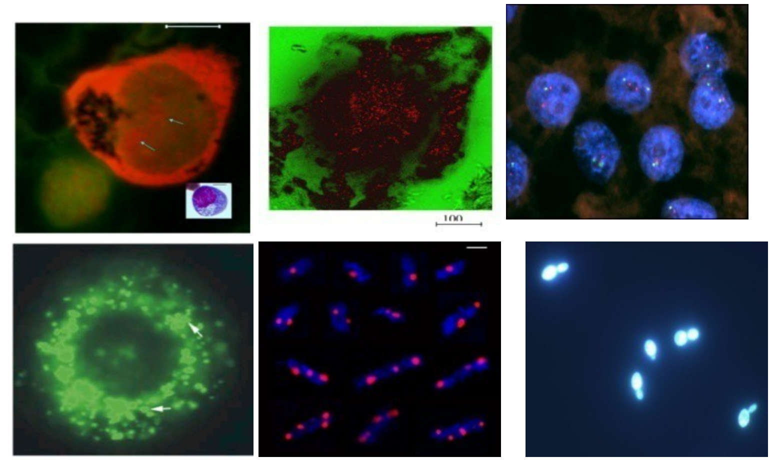




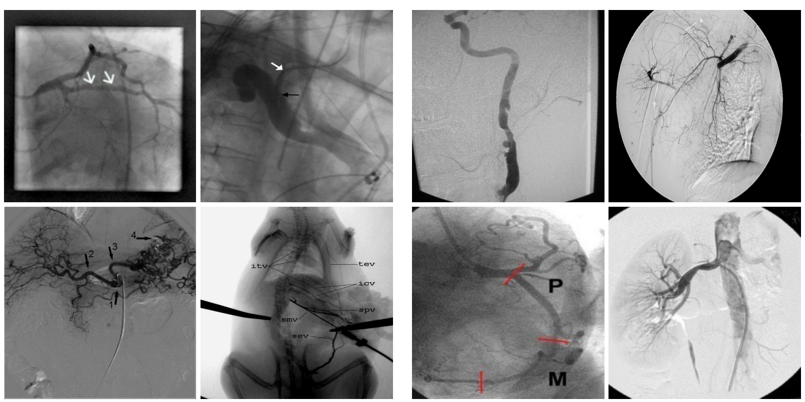
Examples of Correctly Labeled Images



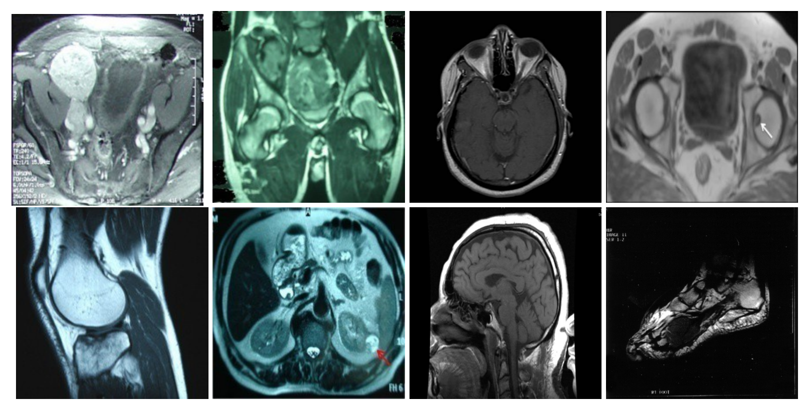
3D Reconstruction



Fluorescence microscopy



Angiography



Magnetic Resonance

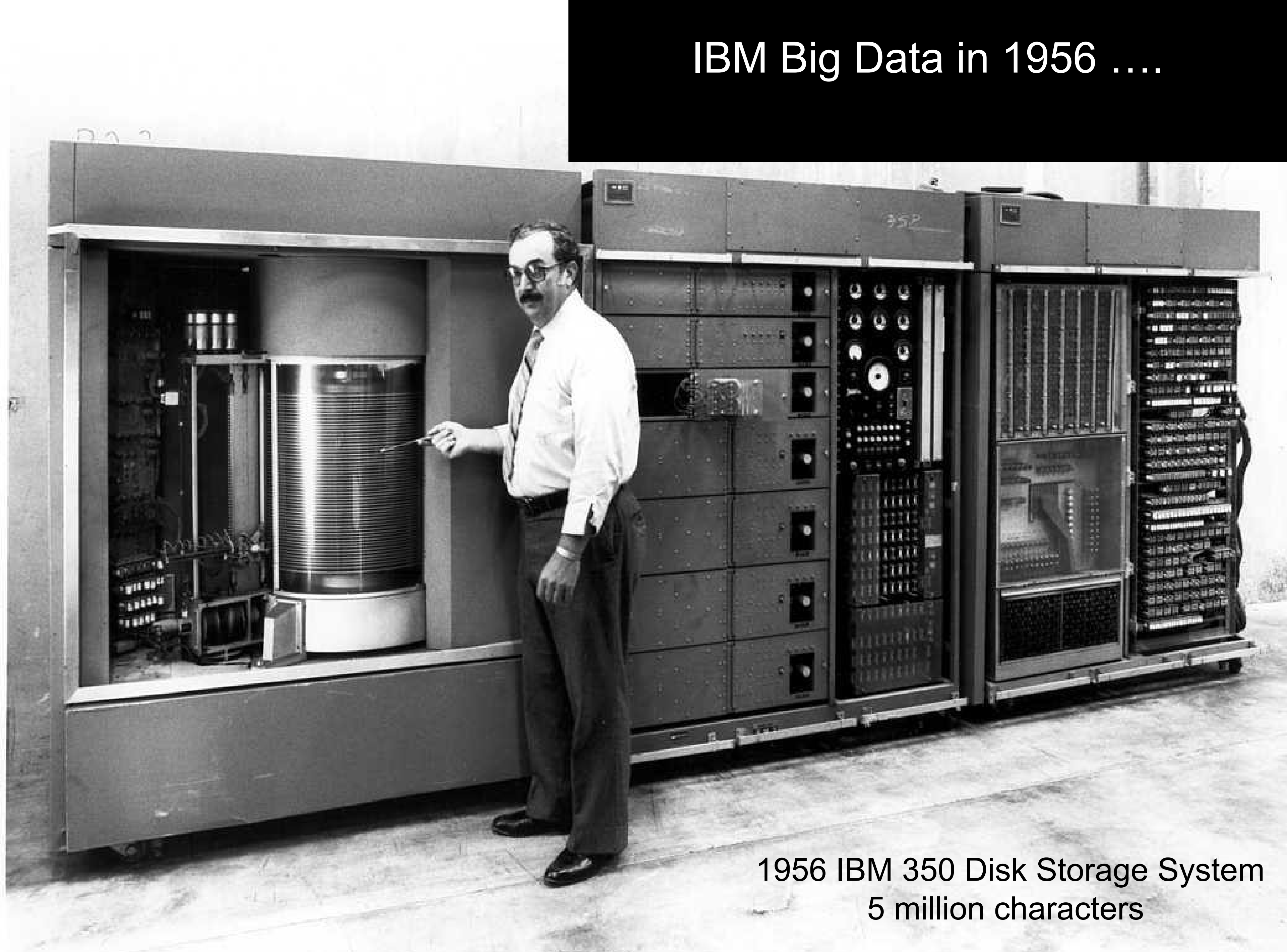
Addressing BIG Problems with BIG Data

Sources and Types of Data

Classifying and Processing Data

A bit about Systems

IBM Big Data in 1956 ….



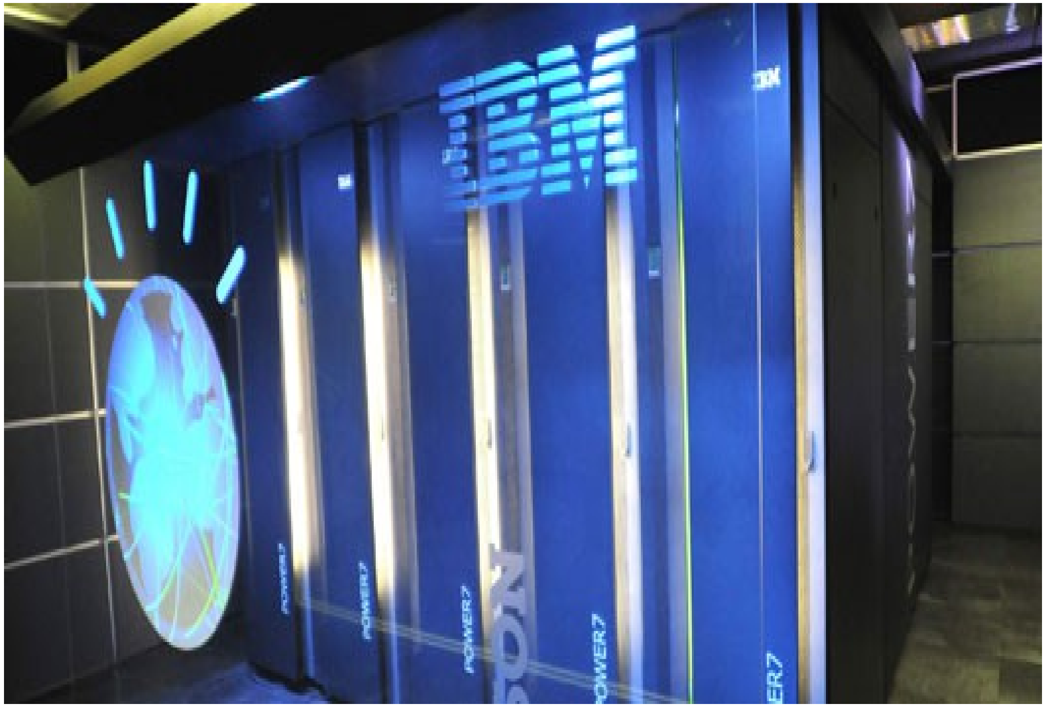
1956 IBM 350 Disk Storage System
5 million characters

Image Src: IBM

Deep Blue

- 1996 / 1997
- 240 Power 2 (p2sc)
  - 30 nodes x 8 chips
- Chess Position Evaluation ASIC
  - 480 (30 x 8 x 2)

Watson



- 360 Power 7 chips
  - 90 nodes x 4 proc.
- 16 TB memory
  - total
- 4 TB disk

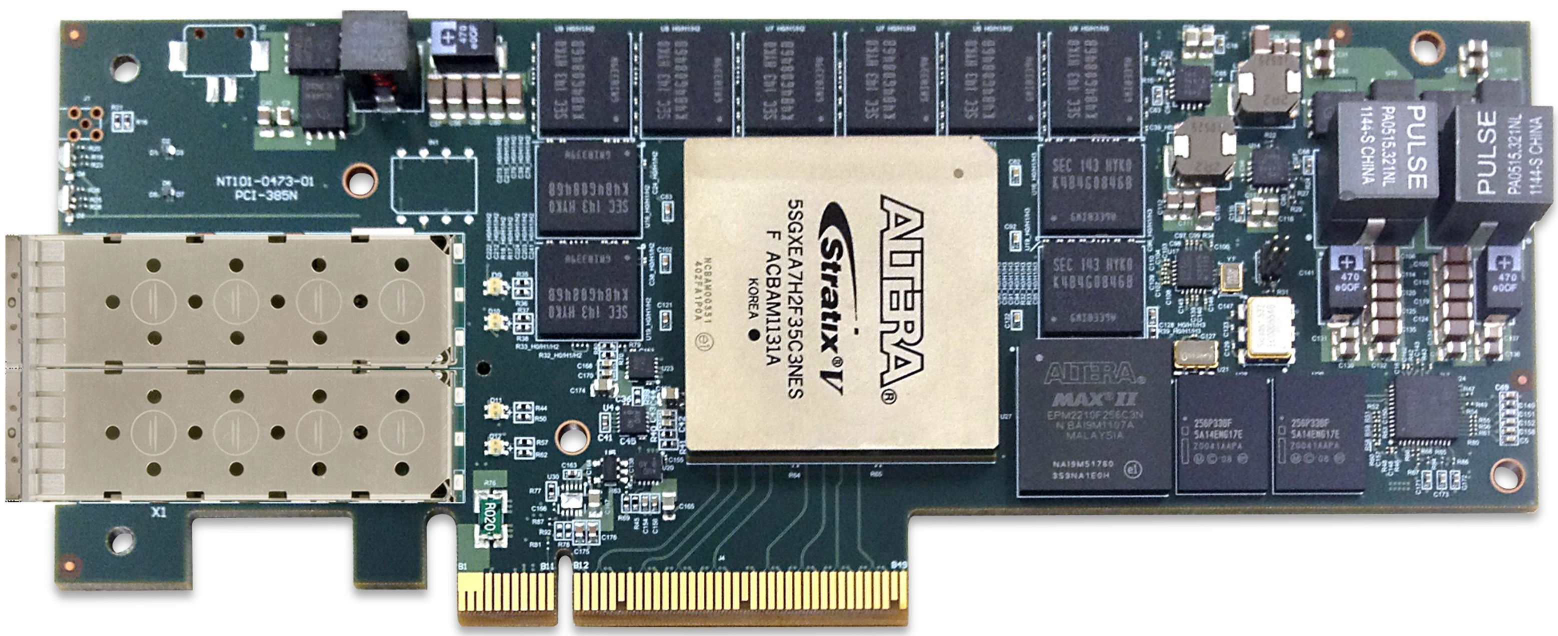
Example Power 7 Big Data System

- **Typical Big Data System**
- **Locally attached storage**
- **Medium strength network (10Gb Ethernet)**
- **16 Power 7 cores per node, 30HDDs**
- **180TB total**
- **< 7 minute Terasort**

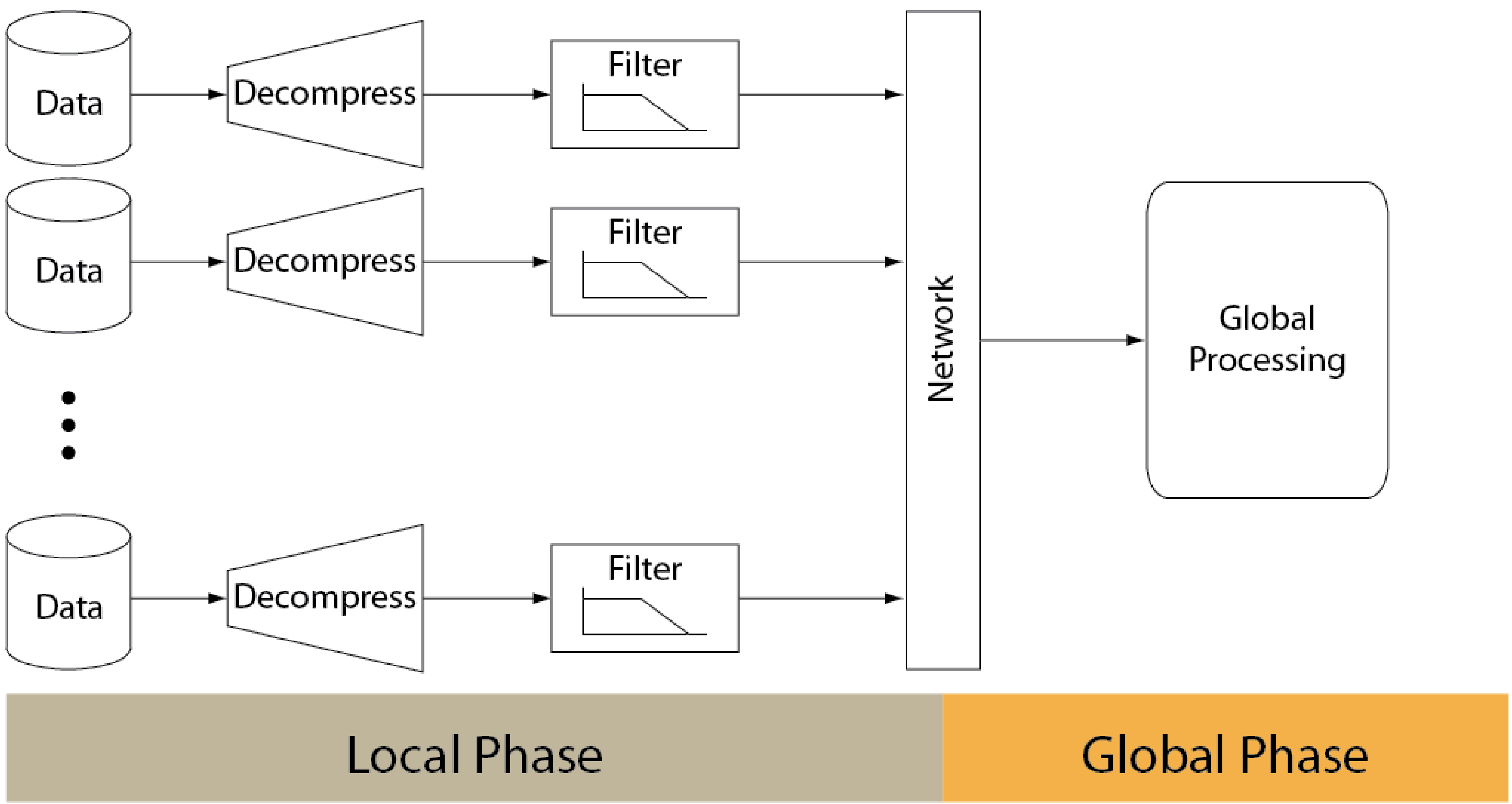
ARL Power Linux Big Data Cluster

48

Accelerating Big Data



Local and Global Compute Phases



High-Performance and Compact Architecture for Regular Expression Matching on FPGA

July 2012 (vol. 61 no. 7)
pp. 1013-1025

Yi-Hua Edward Yang, University of Southern California, Los Angeles
Viktor K. Prasanna, University of Southern California, Los Angeles

DOI Bookmark: http://doi.ieeecomputersociety.org/10.1109/TC.2011.129

We present the design, implementation and evaluation of a high-performance architecture for regular expression matching (REM) on field-programmable gate array (FPGA). Each regular expression (regex) is first parsed into a concise token list representation, then compiled to a modular nondeterministic finite automaton (RE-NFA) using a modified version of the McNaughton-Yamada algorithm. The RE-NFA can be mapped directly onto a compact register-transistor level (RTL) circuit. A number of optimizations are applied to improve the circuit performance: 1) spatial stacking is used to construct an REM circuit processing m\ge 1 input characters per clock cycle; 2) single-character constrained repetitions are matched efficiently by parallel shift-register lookup tables; 3) complex character classes are matched by a BRAM-based classifier shared across regexes; 4) a multipipeline architecture is used to organize a large number of RE-NFAs into priority groups to limit the I/O size of the circuit. We implemented 2,630 unique PCRE regexes from Snort rules (February 2010) in the proposed REM architecture. Based on the place-and-route results from Xilinx ISE 11.1 targeting Virtex5 LX-220 FPGAs, the proposed REM architecture achieved up to **11 Gbps** concurrent throughput for various regex sets and up to 2.67x the throughput efficiency of other state-of-the-art designs.

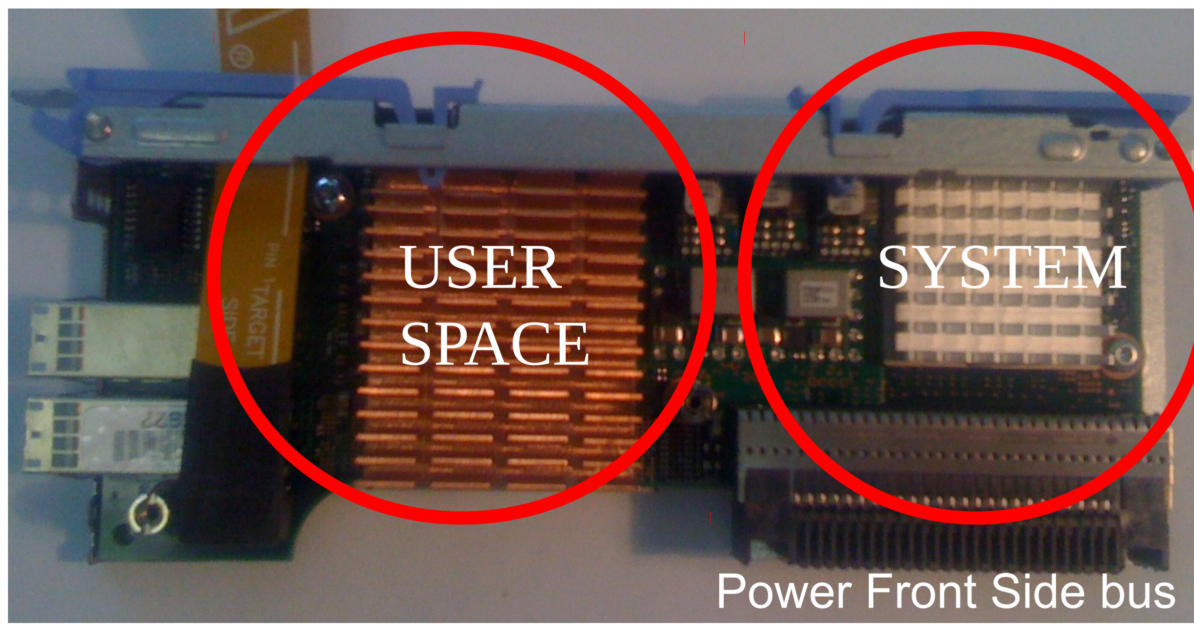
FPGA Accelerated Inflate Prototype Hardware Measurements

- Reference is standard gzip on 3.6GHz Power (linux)
- Benchmarks taken from the Canterbury corpus and the Large corpus (http://corpus.canterbury.ac.nz/descriptions/)

File	Original File Size	Compressed File Size	Compress Ratio	zlib Inflate Time (ms)	FPGA Inflate Time (ms)	FPGA Throughput	FPGA Speedup
asyoulik.txt	**125179**	**59320**	**2.11**	**0.59**	**0.040**	**3.16 GB/s**	**14.99X**
alice29.txt	**152089**	**65188**	**2.33**	**0.67**	**0.047**	**3.21 GB/s**	**14.06X**
lcet10.txt	**426754**	**172770**	**2.47**	**1.77**	**0.129**	**3.31 GB/s**	**13.75X**
plrabn12.txt	**481861**	**241161**	**2.0**	**2.33**	**0.147**	**3.28 GB/s**	**15.87X**
ptt5	**513216**	**67529**	**7.6**	**1.69**	**0.150**	**3.42 GB/s**	**11.27X**
pi.txt	**1000000**	**662248**	**1.51**	**4.67**	**0.334**	**2.99 GB/s**	**13.98X**
world192.txt	**2473400**	**840481**	**2.94**	**9.01**	**0.739**	**3.35 GB/s**	**12.20X**
bible.txt	**4047392**	**1421406**	**2.85**	**15.17**	**1.205**	**3.36 GB/s**	**12.58X**
e.coli	**4638690**	**1891365**	**2.45**	**16.76**	**1.391**	**3.33 GB/s**	**12.04X**

D. Jamsek, A. Martin, K. Agarwal

Accelerated Shared Memory Power Linux Research Prototype



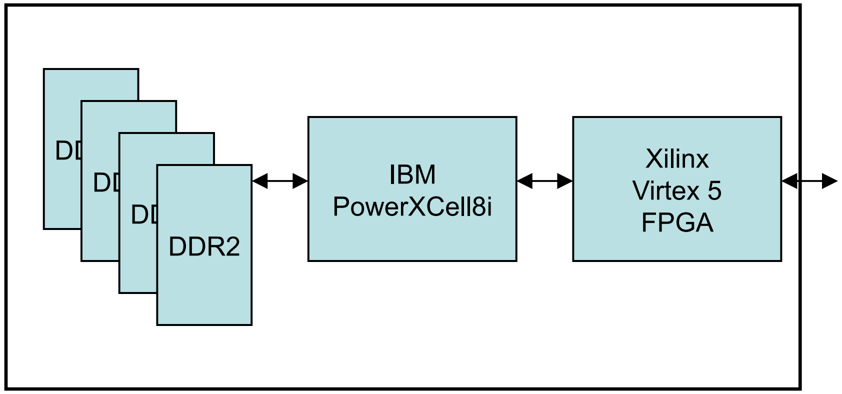
Advantages of Shared Memory Accelerator

- **Virtual Addressing**
  - Removes the requirement for pinning system memory for PCIe transfers
    - Eliminates the copying of data into and out of the pinned DMA buffers
    - Eliminates the operating system call overhead to pin memory for DMA
  - Accelerator can work with same addresses that the processors use
    - Pointers can be de-referenced same as the host application
      - Example: Enables the ability to traverse data structures

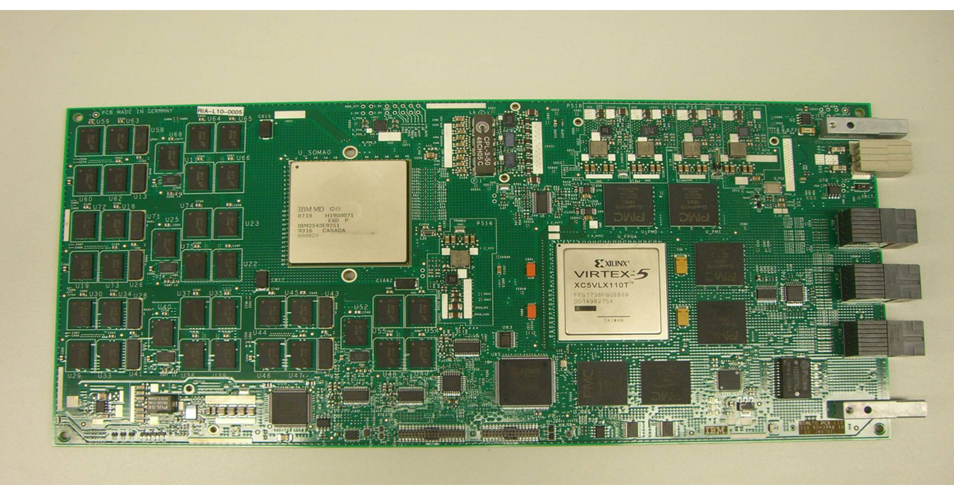
- **Elimination of Device Driver**
  - Direct communication with Application
  - No requirement to call an OS device driver or Hypervisor function for mainline processing

- **Enables Accelerator Features not possible with PCIe**
  - Enables efficient Hybrid Applications
    - Applications partially implemented in the accelerator and partially on the host CPU
  - Visibility to full system memory
  - Simpler programming model for Application Modules

QPACE PowerXCell8i node card and system.



QPACE node card.



Linear Algebra Processor (UT Austin)

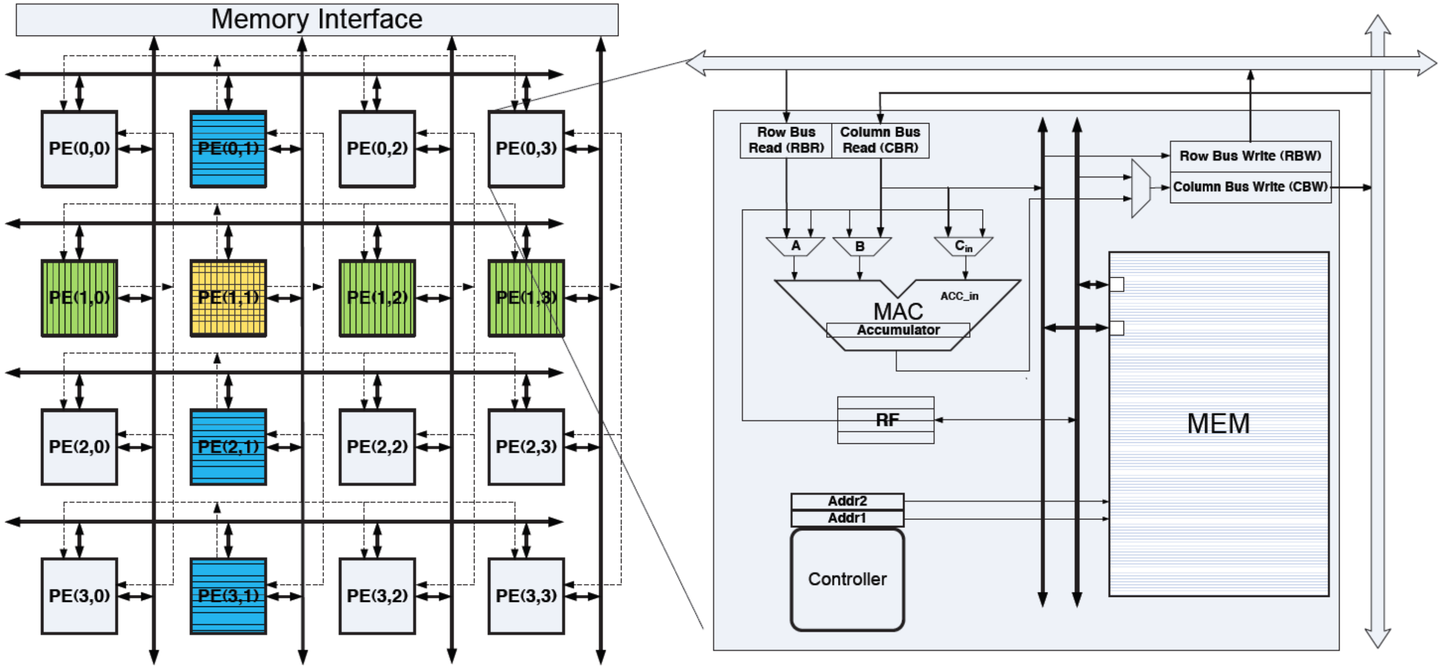


Figure 1: LAP Architecture. The highlighted PEs on the left illustrate the PEs that own the current column of $4 \times k_c$ matrix A and the current row of $k_c \times 4$ matrix B for the second rank-1 update ($p = 1$). It is illustrated how the roots (the PEs in second columns and row) write elements of A and B to the buses and the other PEs read these.

A. Pedram, e.a.

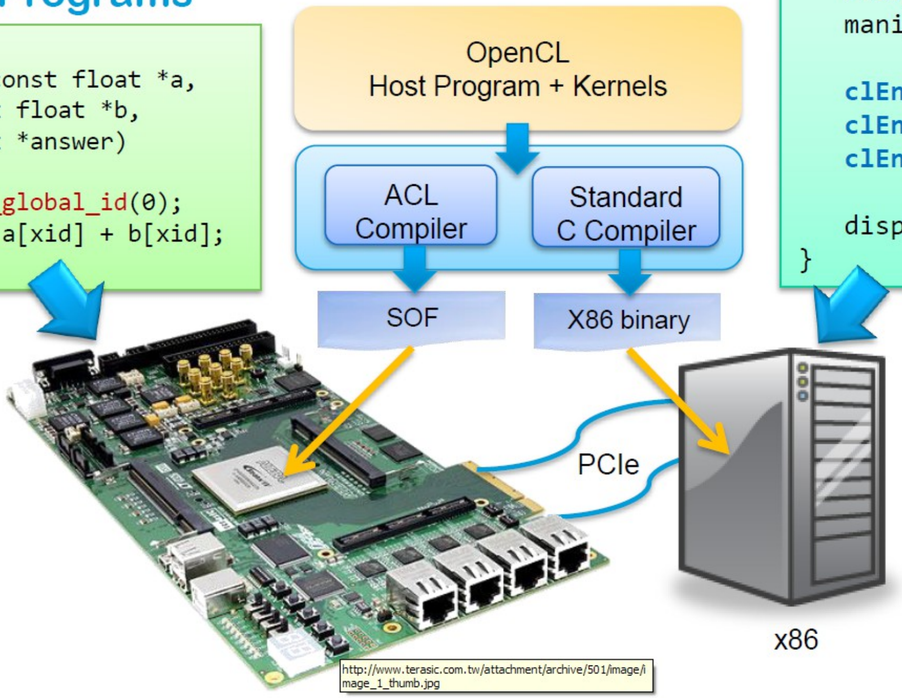
Compiling OpenCL to FPGAs

Kernel Programs

```
__kernel void
sum(__global const float *a,
__global const float *b,
__global float *answer)
{
int xid = get_global_id(0);
answer[xid] = a[xid] + b[xid];
}
```

Host Program

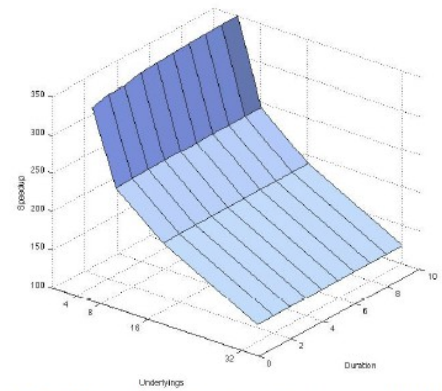
```
main()
{
   read_data_from_file( … );
   manipulate_data( … );

   clEnqueueWriteBuffer( … );
   clEnqueueKernel(…, sum, …);
   clEnqueueReadBuffer( … );

   display_result_to_user( … );
}
```



IBM Wall Street Center of Excellent Demo

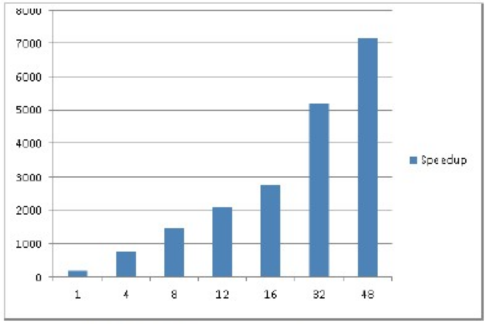
#2: CHREC Multi Asset Barrier Option Pricing Heston Model

- CHREC used UBS model
  - Hand coded VHDL
  - Achieved 100X+ versus Intel core
  - Converted to 32b integer



Single FPGA Speedup : Speedup vs. Duration vs. Underlyings

- OpenCL Double Precision: ~20X+
  - Versus 3GHz Sandy Bridge cores
  - 8 to 64 underlying assets
  - Nallatech PCIe-385N, 23W!



Multiple FPGA's, 10 years, 16 assets

Note *: Stochastic Differential Equations

C-based SSE2 optimized baseline on Intel Sandy Bridge E5-2687 core at 3.1 GHz; Stratix IV E530 FPGA at 125 MHz

http://www.chrec.org/pubs/Sridharan_SAAHPC12.pdf

Vision for seamless acceleration: C on KOVA prototype

- Target C/C++/Fortran loops for acceleration with minimal code changes
- Compiler generates VHDL implementing loop and ELF binary that replaces loop w. FPGA call
- Compiler generated shared-memory communication between CPU/FPGA (no user-visible API)
- Basic implementation running on Austin system: Matrix-matrix multiply, Levenshtein-distance

```
M[0][0] = 0;
for (i = 1; i <= Ni; i++)
  M[i][0] = -i;
for (j = 1; j <= Nj; j++)
  M[0][j] = -j;
for (i = 0; i < NMAX; i++)
  STR1[i] = rand_char();
  STR2[i] = rand_char();
```

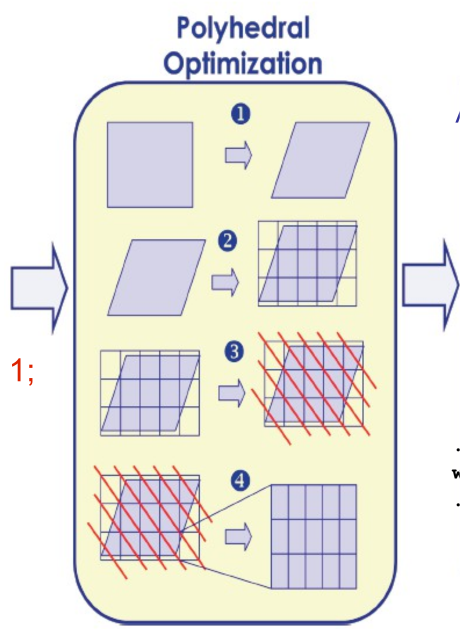
initialize

```
for(i = 1; i <= Ni; i++) {
  for(j = 1; j <= Nj; j++) {
    short score = (STR1[i-1] == STR2[j-1]) ? 0 : 1;
    short a = min (M[i][j-1], M[i-1][j]) + 1;
    short b = min (a, M[i-1][j-1] + score);

    M[i][j] = b;
  }
}
```

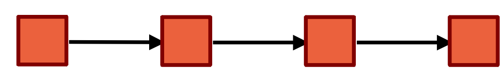
accelerate on FPGA

```
for(i = 1; i <= Ni; i++)
  for(j = 1; j <= Nj; j++)
    printf ("%d\n", M[i][j]);
```

print result




```
M[0][0] = 0;
for (i = 1; i <= Ni; i++)
  M[i][0] = -i;
for (j = 1; j <= Nj; j++)
  M[0][j] = -j;
for (i = 0; i < NMAX; i++)
  STR1[i] = rand_char();
  STR2[i] = rand_char();

call_hardware (&STR1, &STR2, &M);
```



```
for(i = 1; i <= Ni; i++)
  for(j = 1; j <= Nj; j++)
    printf ("%d\n", M[i][j]);
```

A. Jacob

Example Self-Accelerating Adaptive Processor

"LegUp.org" High-Level Synthesis

LegUp accepts a standard C program as input and automatically compiles the program to a hybrid architecture containing an FPGA-based MIPS soft processor and custom hardware accelerators that communicate through a standard bus interface. In the hybrid processor/accelerator architecture, program segments that are unsuitable for hardware implementation can execute in software on the processor.

"Our long-term vision is to fully automate the flow in Fig. 1, thereby creating a self-accelerating adaptive processor in which profiling, hardware synthesis and acceleration happen transparently without user awareness."

University of Toronto / Altera
http://www.eecg.toronto.edu/~janders/fpga60-legup.pdf

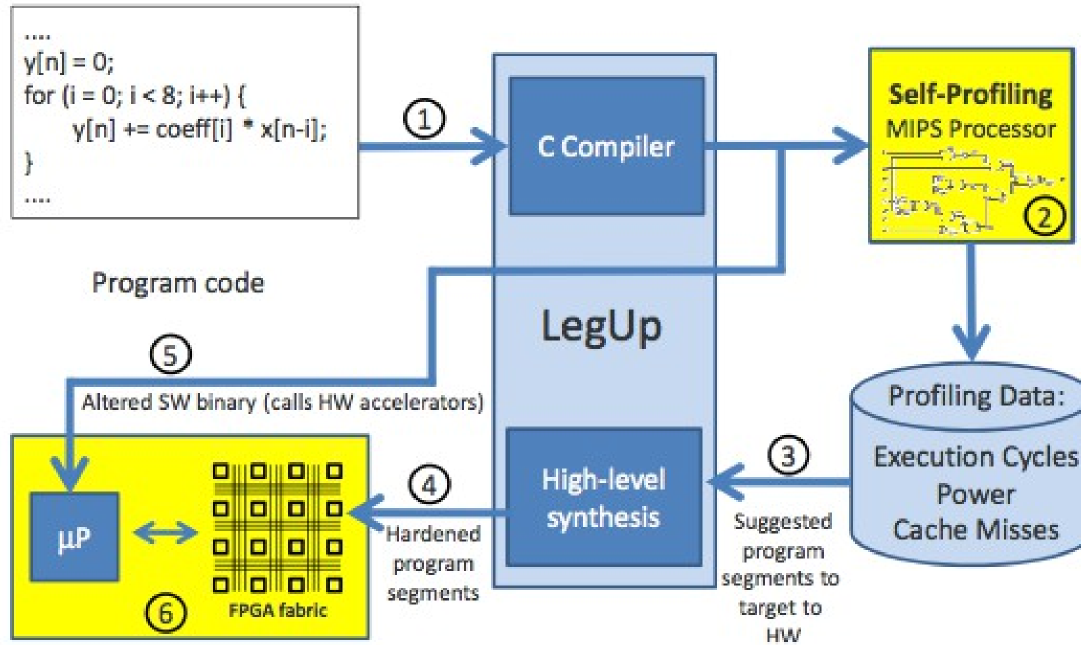


Figure 1: Design flow with LegUp.

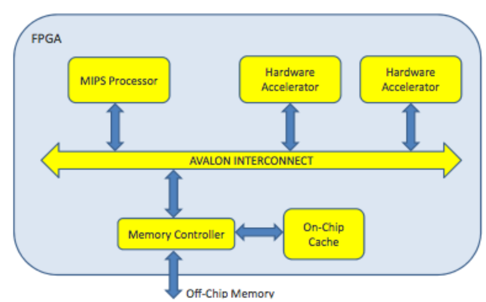


Figure 2: Target system architecture.

- There is no shortage of Big Problems that require Big Data
- The Nature of Data in IT is changing.
- Volume – Data doubling every two years
- Variety – Heading to a trillion devices; Unstructured data;
- Image/Video linked to Mobile & Social growth
- Velocity – Sometimes all you have is milliseconds to respond
- Veracity – My business, finances, safety, health, life depend on
- the quality of the data … and I'm not sure I can trust what I got
- Machines (in part) got us into this flood of Big Data. We need machines to help us out by:
-

1)Finding and exploiting the explicit and implicit structure in Unstructured Text and Image Data

2)Raising the level of interaction by presenting Natural Language interface to sophisticated analytics

THINK

BIG