

# Big Data for Central Banking

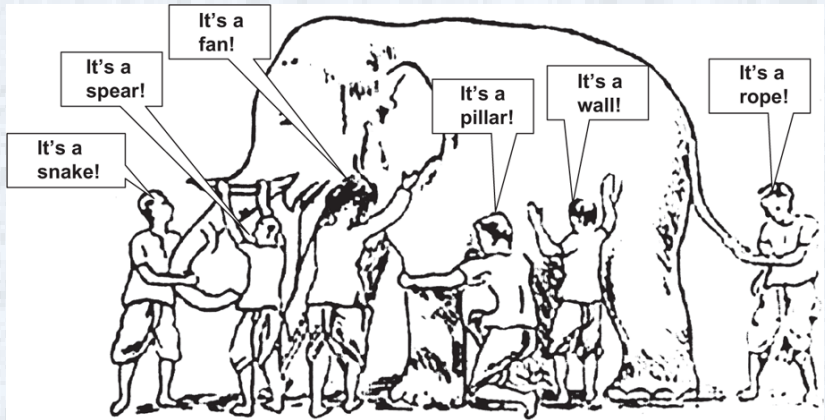
शिवकुमार G. Sivakumar சிவகுமார்

Computer Science and Engineering  
भारतीय प्रौद्योगिकी संस्थान मुंबई (IIT Bombay)  
siva@iitb.ac.in

- Data Scientist's Dream: AI/ML meets Web 3.0 & SMAC + IoT
- Data Engineer's Job: Big Data Frameworks- Hadoop, Spark, Flink
- Data Governance's Nightmare: Quality, Security, Privacy, Rol
- RBI's CIMS (Centralized Information Management System)



# One Single Truth? अन्ध-गज न्यायः



**Note:** The risks of analytical thinking and fragmentation of knowledge



## Aug 24, 2018. New York Fed Staff Nowcast

- The New York Fed Staff Nowcast for 2018:Q3 stands at 2.0%.
- News from this week's data releases decreased the nowcast for 2018:Q3 by 0.4 percentage point.
- The negative impact of higher than expected manufacturers' inventories drove most of the decrease.

2018:Q3 | 2018:Q2 | 2018:Q1 | 2017:Q4

Last Release 11:15am EST Aug 24, 2018

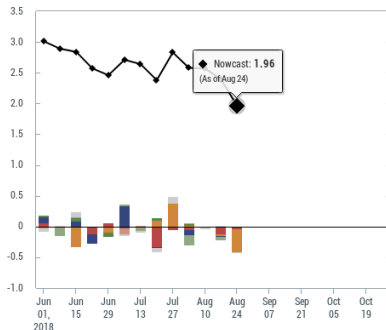
ARCHIVE

LAYOUT

◆ The New York Fed Staff Nowcast ○ Advance GDP estimate □ Latest GDP estimate

■ Housing and construction ■ Manufacturing ■ Surveys ■ Retail and consumption ■ Income ■ Labor ■ International trade ■ Others

Percent (annual rate)



Expand

### Data Flow (Aug 24, 2018)

Model Update	Release Date	Data Series	Actual	Impact	Nowcast GDP Growth
Aug 24					1.96
	8:30AM Aug 24	Manufacturers' inventories: Durable goods	1.25	-0.32	
	8:30AM Aug 24	Manufacturers' unfilled orders: All industries	0.01	0.01	
	8:30AM Aug 24	Manufacturer's shipments: Durable goods	-0.21	-0.02	
	8:30AM Aug 24	Manufacturers' new orders: Durable goods	-1.69	-0.05	
	10:00AM Aug 23	New single-family houses sold	-1.72	-0.03	
		Data revisions		-0.01	
Aug 17					2.39



शिवकुमार

G. Sivakumar சிவகுமார் Computer Science and Engineering

भारतीय प्रौद्योगिकी संस्थान मुंबई (IIT Bombay)

Big Data for Central Banking

STAFF REPORTS

## Macroeconomic Nowcasting and Forecasting with Big Data

November 2017 Number 830

*JEL classification:* C32, C53, C55, E32

Authors: Brandyn Bok, Daniele Caratelli, [Domenico Giannone](#), [Argia M. Sbordone](#), and [Andrea Tambalotti](#)

Data, data, data . . . Economists know it well, especially when it comes to monitoring macroeconomic conditions—the basis for making informed economic and policy decisions. Handling large and complex data sets was a challenge that macroeconomists engaged in real-time analysis faced long before “big data” became pervasive in other disciplines. We review how methods for tracking economic conditions using big data have evolved over time and explain how econometric techniques have advanced to mimic and automate the best practices of forecasters on trading desks, at central banks, and in other market-monitoring roles. We present in detail the methodology underlying the New York Fed Staff Nowcast, which employs these innovative techniques to produce early estimates of GDP growth, synthesizing a wide range of macroeconomic data as they become available.

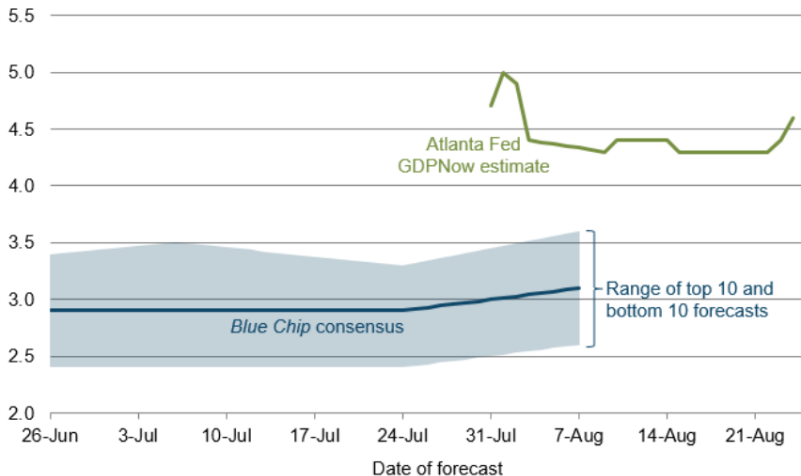
[Available only in PDF](#)



# Evolution of Atlanta Fed GDPNow real GDP estimate for 2018: Q3

Quarterly percent change (SAAR)

**GDPNow™**

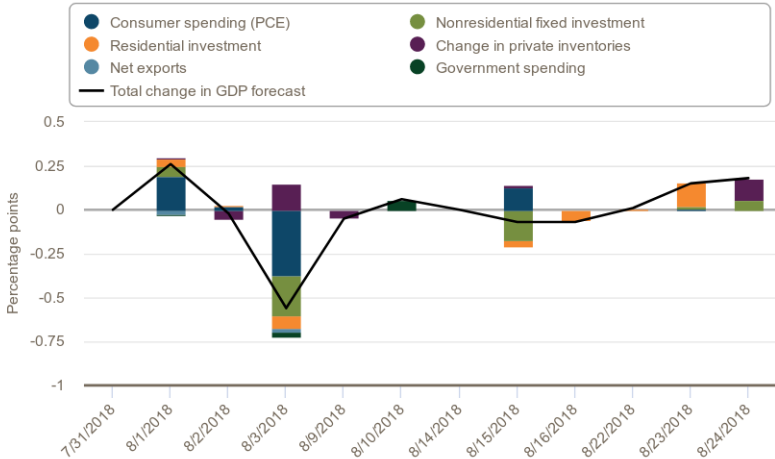


Sources: *Blue Chip Economic Indicators* and *Blue Chip Financial Forecasts*

Note: The top (bottom) 10 forecast is an average of the highest (lowest) 10 forecasts in the *Blue Chip* survey.



## Changes in subcomponent contributions to GDPNow real GDP growth forecasts ≡ Export



Source: Atlanta Fed



FEDERAL RESERVE BANK of ATLANTA



## Atlanta Fed GDPNow estimates for 2018: Q2. growth rates and changes

Date	Major Releases	GDP	PCE	Equip- ment	Intell. prop. prod.	Nonres. struct.	Resid. inves.	Govt.	Exports	Imports	Change in net exp.	Change in CIPI
30-Apr	Initial nowcast	4.1	3.3	9.1	4.4	2.9	3.5	0.5	6.1	3.6	7	26
	Adv. Econ. Indicators/GDP (5/30),											
31-May	Personal inc., NIPA underlying detail	4.7	3.4	5.4	6.4	5.6	0.8	0.6	7.3	2.9	19	43
4-Jun	M3 Manufacturing, Auto sales	4.5	3.3	4.3	6.4	7.2	1.8	0.4	7.5	3.2	18	42
5-Jun	ISM Nonmanufacturing Index	4.6	3.3	4.4	6.4	7.2	1.9	0.4	7.5	3.2	18	42
6-Jun	International trade	4.5	3.4	4.0	6.4	7.2	1.9	0.4	7.7	4.0	13	42
8-Jun	Wholesale trade	4.6	3.4	4.0	6.4	7.2	1.9	0.4	7.7	4.1	13	45
12-Jun	CPI, Monthly Treasury Statement	4.6	3.4	4.0	6.4	7.2	1.6	0.8	7.7	4.0	13	45
13-Jun	Producer Price Index	4.6	3.4	4.0	6.4	7.3	1.4	0.9	7.7	4.0	13	44
14-Jun	Retail trade, Import/Export prices	4.8	3.6	3.9	6.4	7.3	1.4	0.9	7.8	3.8	15	44
15-Jun	Industrial production	4.6	3.7	2.9	6.4	7.4	0.3	0.8	7.1	3.1	17	37
19-Jun	Housing starts	4.7	3.7	2.9	6.4	7.5	2.9	0.8	7.1	3.1	17	37
20-Jun	Existing-home sales	4.7	3.7	2.9	6.4	7.5	2.2	0.8	7.1	3.1	17	37
25-Jun	New-home sales/prices/costs	4.6	3.7	2.9	6.4	7.6	0.6	0.9	7.1	3.1	17	37
	Advance durable manufacturing,											
27-Jun	Advance Economic Indicators	4.5	3.7	2.9	6.4	7.6	0.6	0.9	7.9	1.9	29	20
	GDP (6/28), Personal income and											
29-Jun	outlays, NPA underlying detail tables	3.8	2.7	4.0	6.9	7.6	0.6	0.8	8.0	2.2	28	19
2-Jul	ISM Manuf., Construction spending	4.1	2.9	4.8	6.9	5.3	0.9	1.6	8.4	2.6	27	19

### Maximum forecast of real GDP growth

Employment sit., ISM Manufacturing,

1-Jun	Construction spending	4.8	3.5	6.1	6.4	7.2	1.8	0.4	7.5	3.2	18	43
-------	-----------------------	-----	-----	-----	-----	-----	-----	-----	-----	-----	----	----

### Minimum forecast of real GDP growth

GDP (6/28), Personal income and


29-Jun	outlays, NPA underlying detail tables	3.8	2.7	4.0	6.9	7.6	0.6	0.8	8.0	2.2	28	19
--------	---------------------------------------	-----	-----	-----	-----	-----	-----	-----	-----	-----	----	----

Note: CIPI is "change in private inventories." Changes in net exports and CIPI are both in billions of 2009 dollars (SAAR). All other numbers are quarterly percent changes (SAAR). Table does not necessarily include all estimates for the quarter; see tab "TrackingHistory" in the [online excel file](#) for the entire history.



FEDERAL  
RESERVE  
BANK  
of ATLANTA




**data.gov.in**  
 Open Government Data (OGD) Platform India


Type search keyword

LOGIN | REGISTER

ANALYTICS

195,529 RESOURCES  
 4,267 CATALOGS  
 111 DEPARTMENTS  
 16.48 M TIMES VIEWED  
 5.87 M TIMES DOWNLOADED  
 117 CHIEF DATA OFFICERS  
 6,755 APIs  
 1,427 VISUALIZATIONS

CATALOG



Health Management Information System


INDICATOR DASHBOARD

Education	Demography
Economy	Industries
Transport	Labour And Employment


SUGGEST A DATASET

Your Name

VISUALIZATION GALLERY



SECTORS



Catalogs  
Datasets  
API  
Visual Access  
Services



शिवकुमार

G. Sivakumar

சிவகுமார்

Computer Science and Engineering

भारतीय प्रौद्योगिकी संस्थान मुंबई (IIT Bombay)

Big Data for Central Banking



Instance Title : Yearly Percentage Growth in scheduled Cargo Traffic of All Scheduled Indian Airlines from 2007-08 to 2016-17

OGD Instance Data



Visualizations



Data Selection



Visualization Settings

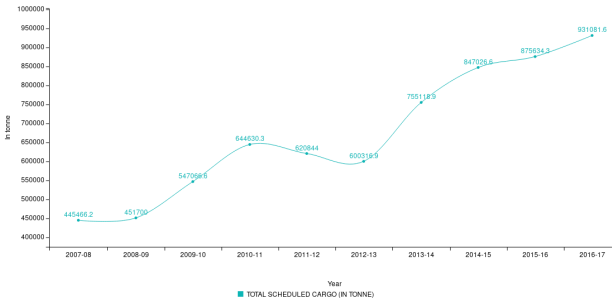


Table



Help Wizard

Cargo Traffic of All Scheduled Indian Airlines from 2007-08 to 2016-17



Some use cases by MoF.

- Railways ticketing data (Internal Migration)
- Satellite data (Urbanization)
- GSTN data (Interstate commerce)
- Is this Big Data?



# Big data and central banking

For central banks, the *flexibility and real-time* availability of big data open up the possibility of extracting more *timely economic signals*, applying new statistical methodologies, enhancing *economic forecasts* and *financial stability* assessments, and obtaining *rapid feedback* on policy impacts.

Economic science, political science, social science, computer science?

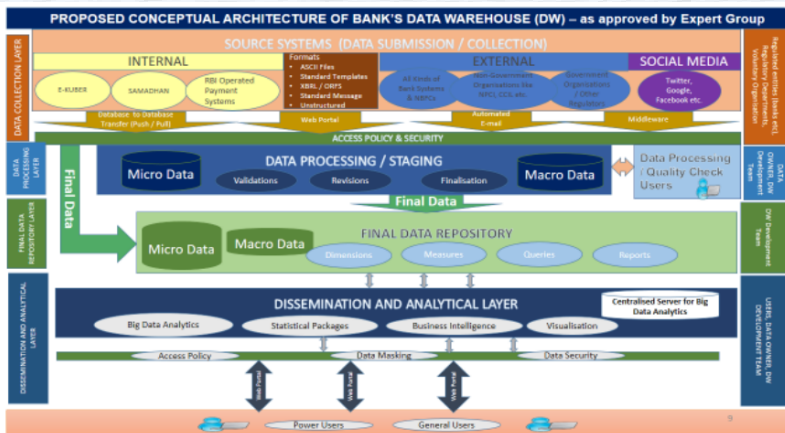


It is critical for a full-service Central Bank, such as the RBI, with diverse responsibilities – *inflation management, currency management, debt management, reserves management, banking regulation and supervision, financial inclusion, financial market intelligence and analysis, and overall financial stability* – to employ **relevant data** and apply the right filters for improving its **forecasting, nowcasting, surveillance and early-warning** detection abilities that all aid **policy formulation**. In the backdrop of ongoing explosion in information gathering, computing capability and analytical toolkits, policy making benefits not only from data collected through **regulatory returns and surveys** but also from large volumes of **structured and unstructured** real-time information sourced from **consumer interactions** in the digital world. Accordingly, it has been decided to gainfully harness the power of Big Data analytics by setting up a Data Sciences Lab within the RBI that will comprise **experts and budding analysts**, internal as well as lateral, who are trained inter alia in Computer Science, Data Analytics, Statistics, Economics, Econometrics and/or Finance. It is envisaged that the unit will become operational by December 2018.

*Apr 2018 RBI press release.*



# RBI's Vision of CIMS



- i) Tier-I: Data Collection Layer (DCL)
- ii) Tier-II: Data Processing Layer (DPL)
- iii) Tier-III: Final Data Repository Layer (FDRL)
- iv) Tier-IV: Dissemination and Analytics Layer (DAL)



# CIMS Major Components

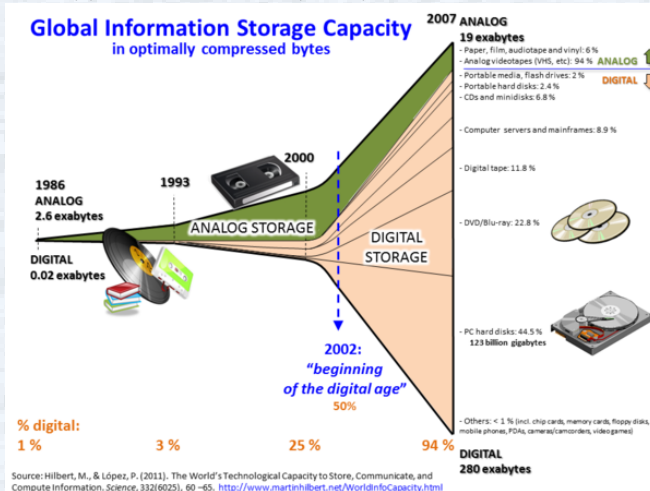
## From draft EoI

- System-to-System *automated data collection* of Macro and Micro data from source systems of banks and other entities to CIMS.
- Staging Area Data Portal (SADP) system for accessing/authorization/ revision of data by banks and entities.
- **Designed Data:** Relational Database Management Systems (RDBMS) based repository with fact and dimension tables.
- **Organic Data:** Data Lake - which is composed of commodity hardware (one time write and multiple reads) as micro level big-data are stored in both structured and unstructured forms.
- Centralised Analytics Server will facilitate **user friendly** analytics and **multi-dimensional view** of data, its trends and relationships/ correlations and the like.



# Data Deluge

[https://en.wikipedia.org/wiki/Big\\_data](https://en.wikipedia.org/wiki/Big_data)



शिवकुमार

G. Sivakumar சிவகுமார் Computer Science and Engineering

भारतीय प्रौद्योगिकी संस्थान मुंबई (IIT Bombay)

Big Data for Central Banking



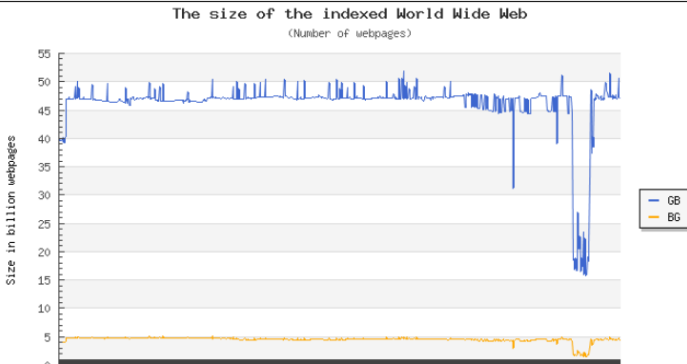
## The size of the World Wide Web (The Internet)

[Tweet](#)

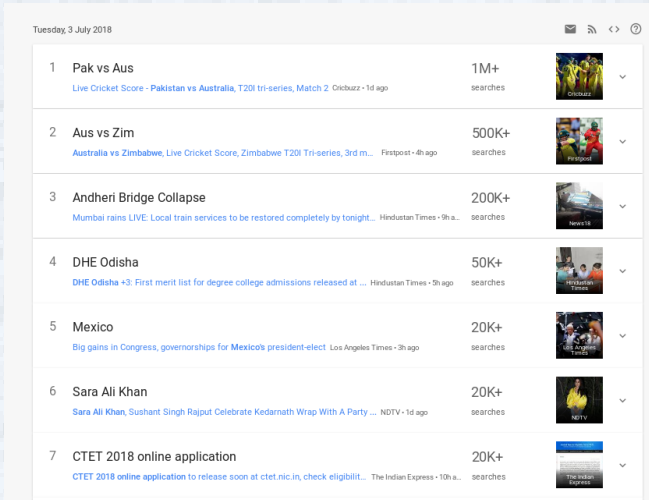
The Indexed Web contains **at least 4.49 billion pages** (Tuesday, 03 July, 2018).

The Dutch Indexed Web contains **at least 154.3 million pages** (Tuesday, 03 July, 2018).

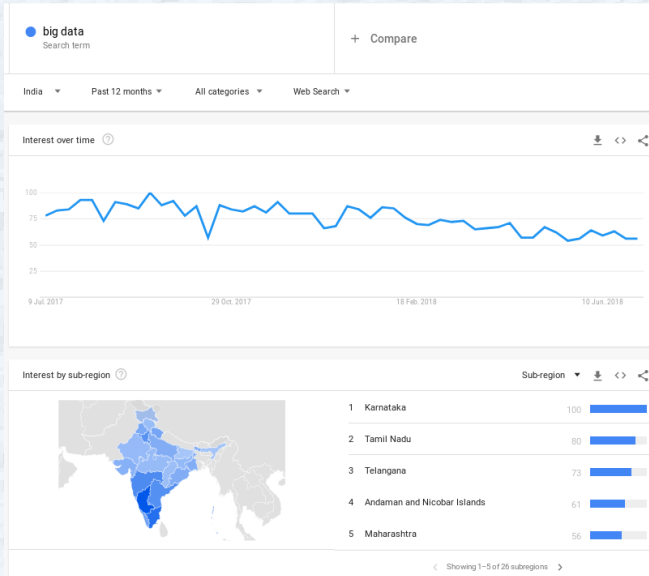
The Indexed Web | [The Dutch Indexed Web](#)

[Last Month](#)[Last Three Months](#)[Last Year](#)[Last Two Years](#)

# Google Search Trends



# Google Topic Search



शिवकुमार

G. Sivakumar

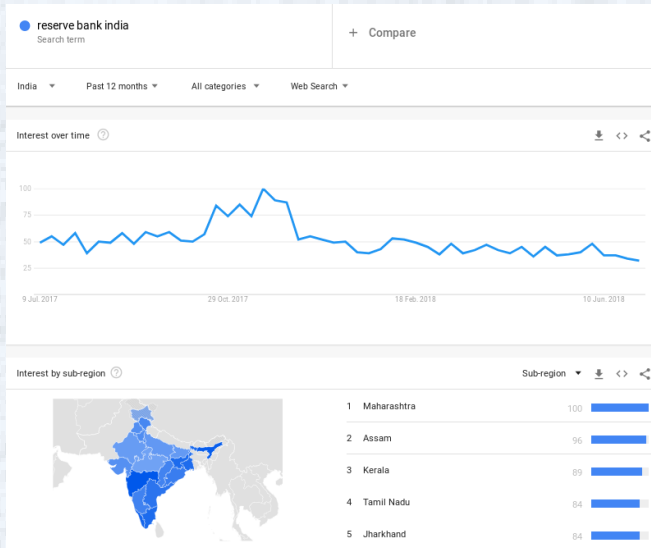
சிவகுமார்

Computer Science and Engineering

भारतीय प्रौद्योगिकी संस्थान मुंबई (IIT Bombay)

Big Data for Central Banking

# Google Topic Search



# Twitter Trends

trends24 India ▾

CURRENT TOP TWITTER TRENDING TOPICS TAG CLOUD FOR INDIA

#HappyOnam #CCVTrailer #LakmeFashionWeek2018  
#HappyRakshaBandhan #IncreaseShowsForMTM #BelgianGP  
#SeemaRajaKaraokeBooth #SaamySquare #LakshmiSuperHit #OthaiyadiPathayila  
#SafetyGuaranteed Rahul Gandhi #Ghoul National Herald #SaturdayMotivation  
#6DaysToGoForAnnanukkuJai AIIMS #WOLMCI #LaLigaOnSony #MamaNahiRojgarChahiye  
#ARSWHU #LIVBHA #GetBehindThem #10YearsOfGourmet #Adangathey Ramlila Maidan  
#NirmalaSitharaman #HappyPhirrBhagJayegi Tajinderpal Singh Toor #MissGrandInIndia #MotherTeresa Rs 46,000  
#6daystogofor60VM #HeroISL Amartya Sen #VaayadiPethaPulla #LakshmiRunningSuccessfully #Siplings #JohnMcCain  
#AtalAsthiKalashYatra #SamvidhanBachao #RajasthanGauravYatra #5daystogofor60VM #Section777 Nearly 1.2 #SarkarWorkingStills  
#FranklySpeakingWithAmarinder Gopal Bose #SillyFellows #LFWWF2018 #Kalki #LetsRideTogether #DibbaLautao #Vijayakanth #MohammadIrfaan  
India of 60s #AdangatheyTrailer #100HoursOfferByCasaGrand #RakshaMatlabRokNahi #Unnao #GlimpseOfPawanKalyanCDP #RahulGandhiPressMeet #SundayMotivation



शिवकुमार

G. Sivakumar

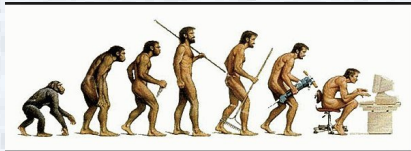
சிவகுமார்

Computer Science and Engineering

भारतीय प्रौद्योगिकी संस्थान मुंबई (IIT Bombay)

Big Data for Central Banking

# Stone Age to Information Age



Homo **Erectus**, Homo **Sapiens**, Homo **Deus** [Yuval  
Noah Harari]

Change due to **Adaptation**  
≠ **Improvement!**

## Technology (Wikipedia Definition)

Technology is the usage and knowledge of tools, techniques, crafts, systems or methods of organization in order to **solve a problem** or **serve some purpose**.

Zero, Wheel, Printing Press, Radio, Lasers, ...

*Any sufficiently advanced technology is indistinguishable from magic.* [Arthur C. Clarke]

- Why **Information Technology** is different?  
Transistor, VLSI, Microprocessor, ...
- **Danger:** Computers are coming! Taking away our jobs!  
**Construction, Farming**, Banking, Surgery, **Composing music, Teaching!**  
Be very scared!



# Web 1.0, Web 2.0, Web 3.0

## Web 1.0 [1990-2005] (*Right to Information*)

- Internet: Info anytime, anywhere, any form
- Like *drinking water from a fire hose*
- Search Engines to the rescue

## Web 2.0 [2005-2015] (*Right to Assembly*)

- Social Networking (Twitter, Facebook, Kolaveri, Flash crowds)
- Producers, not only consumers (Wikipedia, blogs, ...)
- *Proliferated unreliable, contradictory information?*
- *Facilitated malicious uses including loss of privacy, security.*

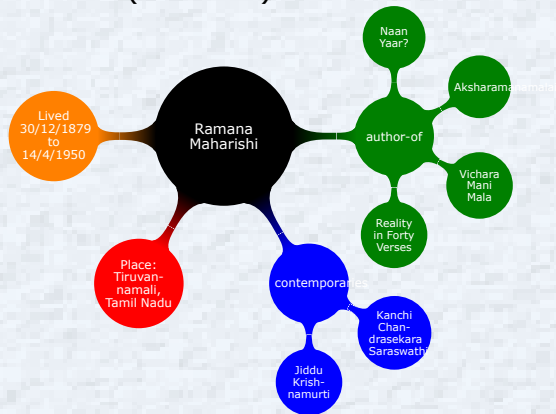
## Web 3.0 [current] (*AI & ML meet Semantic Web*)

- Intelligent Agents that “understand”
- What do you want when you get up and put on computer?
- *I have a dream!*(MLK)



# Semantic Web

The **application layer** tapping the *hardware* (Web 1.0) and OS (Web 2.0)?



Combined, categorized information inferred from various sites, languages. [www.dbpedia.org](http://www.dbpedia.org) comes close today!



# 3rd platform: SMAC + IoT

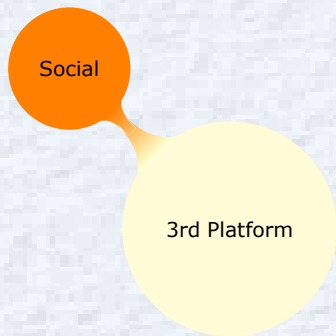
- Main Frame (1960s ...)
- Client Server (1990s ...)
- Today (Handheld, Pervasive Computing)



3rd Platform



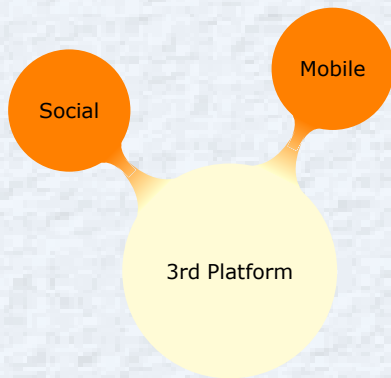
# 3rd platform: SMAC + IoT



- What's App (how many engineers?)
- Facebook, Twitter, GooglePlus ...
- Web 2.0 (Right to Assembly)
- Crowdsourcing (Wikipedia)
- Crowdfunding (no banks!)



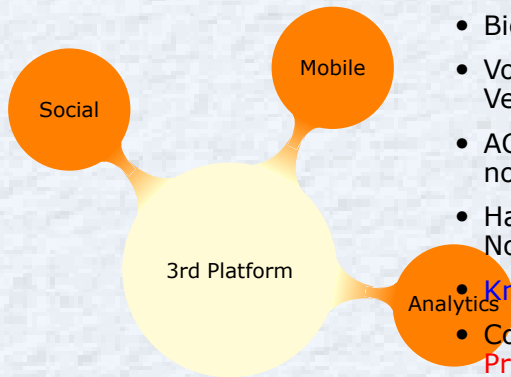
# 3rd platform: SMAC + IoT



- Phone (Smart, Not-so-smart!)
- Wearables! (Google glass, Haptic)
- Internet of "Me" (highly personalized) Business (no *generic* products!)
- **BYOx**: Device security, App/content management nightmare.
- **Data Loss Prevention** (Fortress Approach - Firewall, IDS/IPS - won't work!)



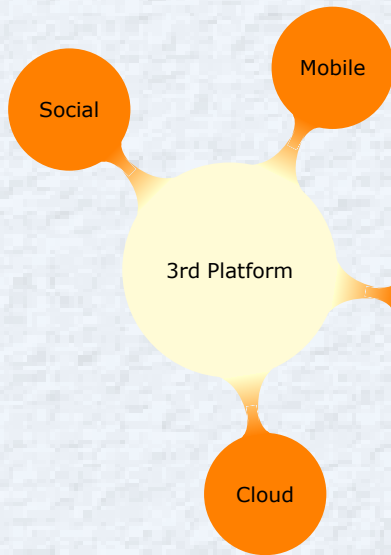
# 3rd platform: SMAC + IoT



- Big Data Analytics
- Volume, Variety, Velocity, Veracity
- ACID properties Database not needed
- Hadoop, Map Reduce, NoSql
- Knowledge is Power!
- Collect, Analyse, Infer, Predict



# 3rd platform: SMAC + IoT



- Moore's law
- What could fit in a building .. room ... pocket ... blood cell!
- **Containers** Analogy from

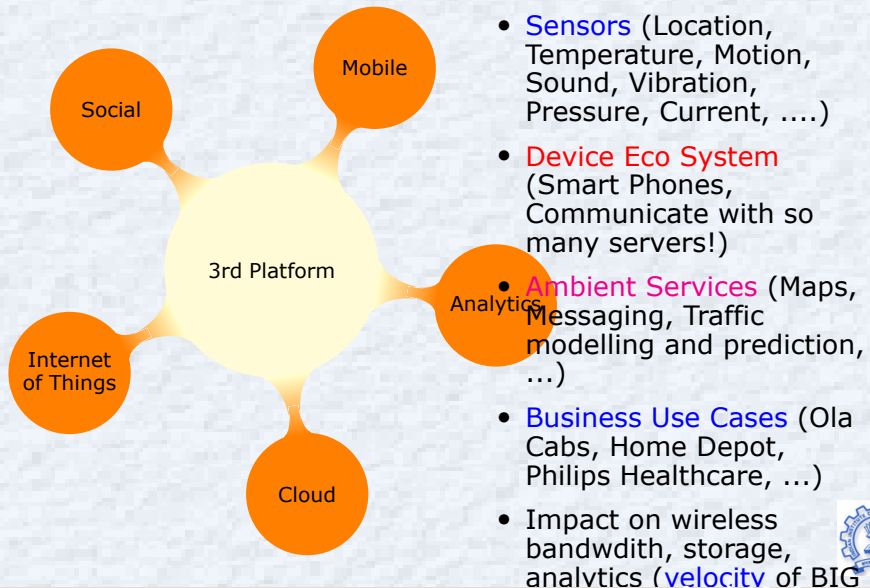
Shipping



- **VMs** separate OS from bare metal (at great cost- Hypervisor, OS image)
- **Docker**- separates apps from OS/infra using containers.
- Like *IaaS, PaaS, SaaS*  
Have you heard of **CaaS**?

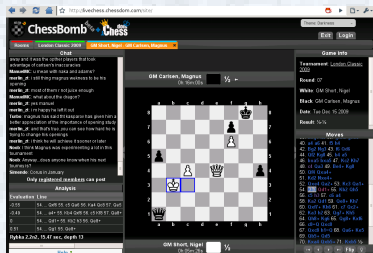


# 3rd platform: SMAC + IoT



# Artificial Intelligence & Machine Learning

- Can AI of computers match NS of humans?
- Old Joke: *Out of sight, out of mind*
- Consider chess, once the *holy grail* of AI. Praggnaananda (prodigy!)



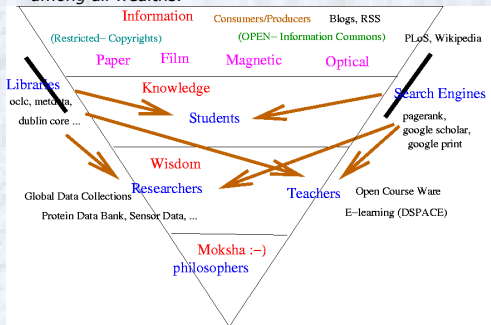
Does not play the human way at all! Mostly parallelized search in hardware (200 million positions/second!)

- December 2017: **AlphaGo Zero** used *reinforcement learning* to teach itself chess in 4 hours! Beat world's best program *Stockfish* comprehensively!



IIT Bombay motto: ज्ञानम् परमम् ध्येयम्  
न चोरहार्यं न च राजहार्यं न भ्रातृभाज्यम् न च भारकारी  
व्यये कृते वर्धत एव नित्यं विद्याधनं सर्वधनप्रधानं

*It cannot be stolen by thieves, cannot be taken away by the king, cannot be divided among brothers and does not cause a load. If spent, it always multiplies. The wealth of knowledge is the greatest among all wealths.*



- How is **learning** affected by the **information deluge**?
- Data is not Information/Knowledge/Wisdom.
- Analogy of Monkeys randomly typing
- How to find Shakespeare's verse?
- Data Mining (My-ning? ;-)
- Has Internet damaged the human brain?

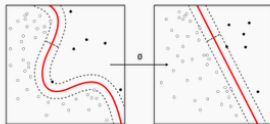


# Data (based) Science

[https://en.wikipedia.org/wiki/Data\\_science](https://en.wikipedia.org/wiki/Data_science)

- Turing award winner Jim Gray terms it the 4th paradigm after Empirical, Theoretical and Computational
- How does Google translate documents?
- Combining data sources to produce new information not contained in any single one!
- How does Alpha-Zero play chess?
- Deep Learning! (Speech Recognition)

## Machine learning and data mining



### Problems

[show]

### Supervised learning

[show]

(classification • regression)

### Clustering

[show]

### Dimensionality reduction

[show]

### Structured prediction

[show]

### Anomaly detection

[show]

### Neural nets

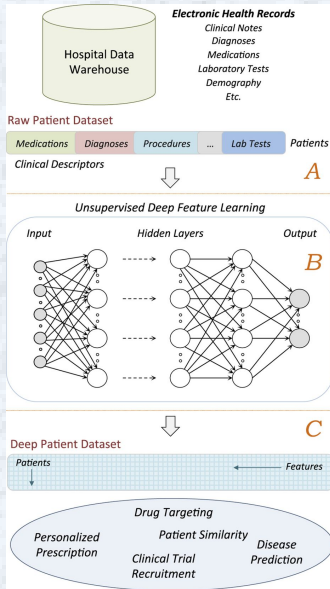
[show]

### Reinforcement learning

[show]



# Deep Patient



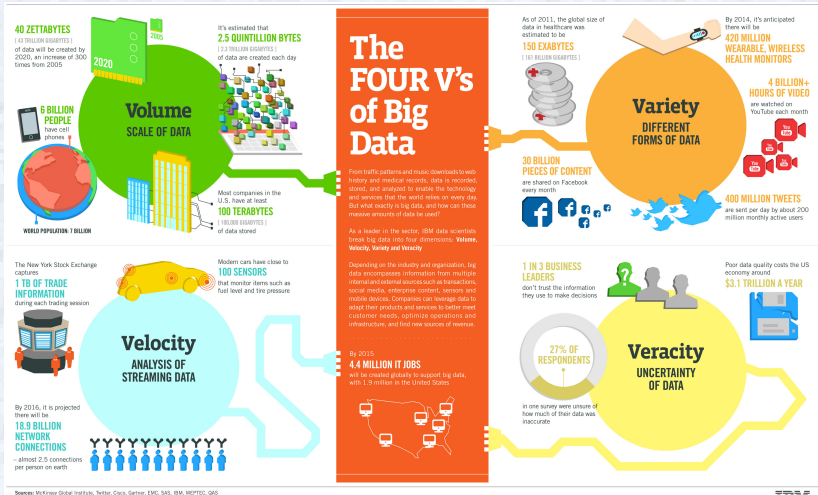
Are doctors practicing **medical science?**

<https://www.nature.com/article>

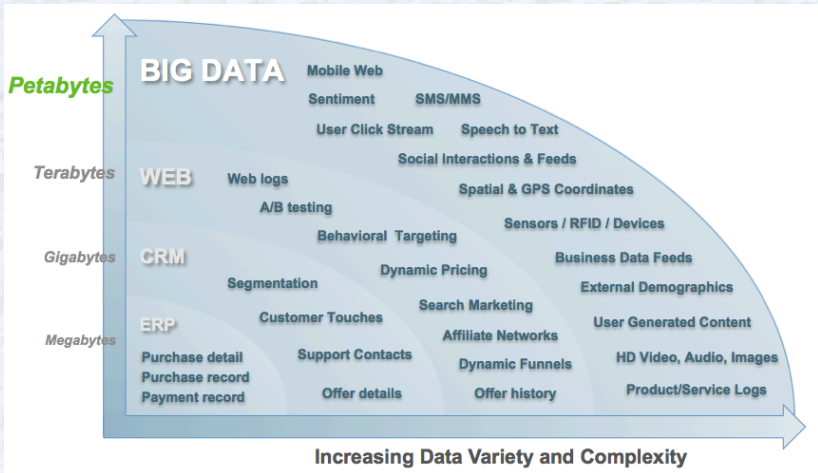
The machine was given no information about how the human body works or how diseases affect us. It found correlations that let it predict the onset of some diseases more accurately than ever, and some diseases, such as schizophrenia, for the first time at all. It does this by creating a vast network of weighted connections that is just too complex for us to understand.



# 4-Vs of Big Data (IBM's Version)



# So, What's Big Data?



# Big Data Technology Stack

- Data Ingestion
  - Extract Transform Load (ETL)
  - continuous or asynchronous, real-time or batched
  - Flume/Kafka
- Data Storage
  - Data Center, Cluster Nodes (Commodity Hardware)
  - HDFS, S3, HBase, Cassandra, MongoDB (document)
- Data Processing/Analytics
  - MapReduce, Spark (in-memory), Storm,
  - R, Python, Java, Scala, ...
  - Mahout (machine learning)
  - Lambda Architecture
- Data Visualization
  - Reporting, Alert, Recommendation, Dashboard
  - Tableau, Kibana

To make all this work, we need *Security, Orchestration*



# Layered Technology Stack



# Why Analytics?

*Opportunity knocks, but once.*

There is a tide in the affairs of men, Which taken at the flood, leads on to fortune. Omitted, all the voyage of their life is bound in shallows and in miseries. *Shakespeare*

*Forewarned is Forearmed.*

चिन्तनीया हि विपदां आदावेव प्रतिक्रिया  
न कूपखननं युक्तं प्रदीप्ते बन्धिना गृहे

The effect of disasters should be thought of beforehand. It is not appropriate to **start digging a well when the house is ablaze with fire.**



# Vishnu ( भूतभव्य भवत्प्रभुः )

॥ हरिः ॐ ॥

विश्वं विष्णुर्वषट्कारो भूत भव्य भवत्प्रभुः ।

- Past (What happened? Reactive)  
Designed Batch/Static Data  
Reports, Standards, Data Harmonization.
- Present (What is happening?)  
Organic Unstructured Streaming/Real-time Data  
Statistical Analysis, Anamolies, Alerts
- Future (What will happen? Pro-active)  
Predictive Forecast, Optimize

Analytics can convert *data to knowledge to wisdom.*



# What's special in Big Data Analytics?

What more/different do we need from

- Databases (RDBMS)  
Used for Transaction processing, consolidation, reports, business performance
- Data Warehouse  
Used to Co-relate various data, strategies for pricing, supply chains, cross-sale, ....

Can these not handle Web log for profiling and recommendations, Sentiment analysis for product evaluation, positioning, marketing?

No. Biggest problem is the **ACID** properties requirement which makes *scale-out* and *fault-tolerance* near impossible.  
Some Weaknesses:

- Disk oriented storage and indexing structures
- Multithreading to hide latency
- Locking-based concurrency control mechanisms
- Log-based recovery



# Big Data Analytics Requirements

Think: Facebook likes, Shopping cart recommendations, ...

- Ingest data at very high speeds and rates
- Scale easily to meet growth and demand peaks
- Support integrated fault tolerance
- Support a wide range of real-time (or “near-real-time”) analytics
- Integrate easily with high volume analytic datastores



# High Speed Data Ingestion

- Support millions of write operations per second at scale
- Read and write latencies below 50 milliseconds
- Do not need ACID-level consistency guarantees (Eventual is fine!)
- Support one or more well-known application interfaces
  - SQL
  - Key/Value
  - Documents (NLP)



# Scaling Requirements

- Scale-out on commodity hardware
- Built-in database partitioning  
Manual sharding and/or add-on solutions not practical Database must automatically implement defined partitioning strategy
- Application should see a single database instance
- Database should encourage scalability best practices  
For example, replication of reference data minimizes need for multi-partition operations



# Data Governance

Most critical, but most neglected (like in this talk!)

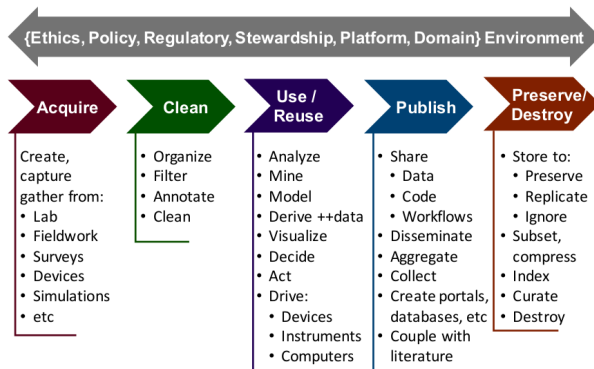


FIGURE 1: The Data Life Cycle and Surrounding Data Ecosystem

*from the National Science Foundation CISE  
AC Data Science Report, October 2016*



Know Thyself (more than the Algorithms know you)

*The algorithms are watching you right now. They are watching where you go, what you buy, who you meet. Soon they will monitor all your steps, all your breaths, all your heartbeats. They are relying on Big Data and machine learning to get to know you better and better. And once these algorithms know you better than you know yourself, they could control and manipulate you, and you won't be able to do much about it. You will live in the matrix ... Yuval Noah Harari*



# What next?

RoI:



"Your recent Amazon purchases, Tweet score and location history makes you 23.5% welcome here."

आचार्यात् पादमादत्ते पादं शिष्यः स्वमेधया ।  
सब्रह्मचारिभ्यः पादं पादं कालक्रमेण च ॥  
one fourth from the  
teacher,  
one fourth from own  
intelligence,  
one fourth from  
classmates,  
and one fourth only with  
time.



## Some useful resources.

- [https://www.bis.org/ifc/events/big\\_data\\_mar2017/bigdata\\_12pres.pdf](https://www.bis.org/ifc/events/big_data_mar2017/bigdata_12pres.pdf)  
(2017 survey of Central Banks on their plans for Big Data)
- [https://www.bis.org/ifc/publ/ifcb44\\_overview\\_rh.pdf](https://www.bis.org/ifc/publ/ifcb44_overview_rh.pdf) (Summary of the proceedings of a recent workshop on Big Data for Central Banks)
- <https://www.newyorkfed.org/research/policy/nowcast>
- <https://www.frbatlanta.org/cqer/research/gdpnow.aspx>
- <https://arxiv.org/pdf/1610.09962.pdf>  
An Experimental Survey on Big Data Frameworks (including Hadoop, Spark, Flink)
- Relevant *Wikipedia* pages

