

Data Analytics, Interpretation and Descriptive Statistics

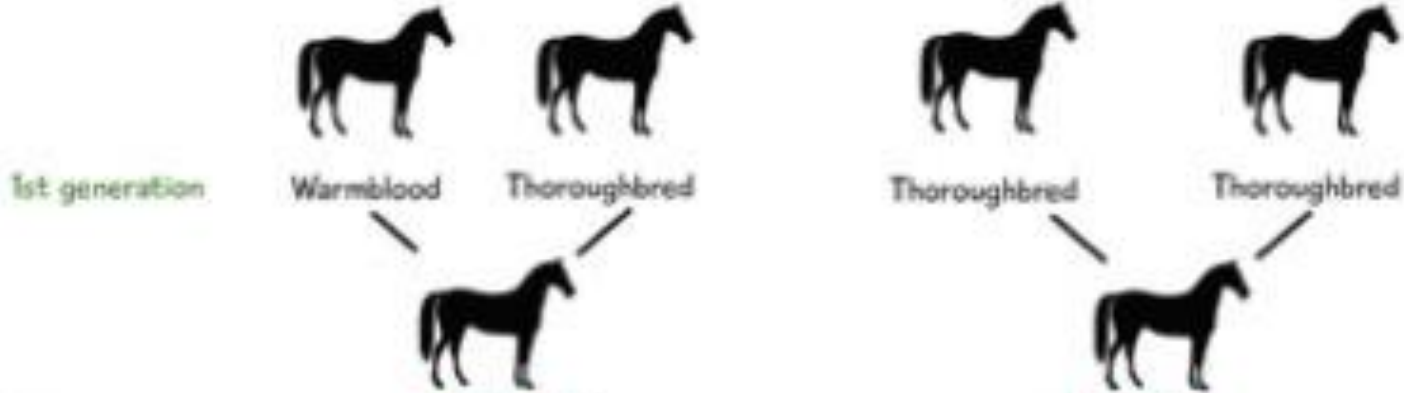
26 Feb 2021 @ MCRHRDI

Sudhir Voleti

Associate Professor of Marketing, ISB

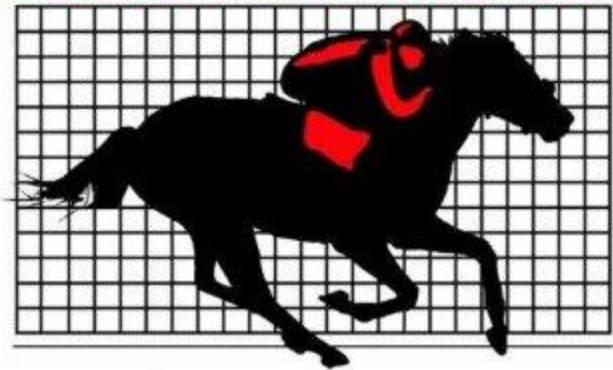
Sudhir_Voleti@isb.edu

Now, let's look at several generations. The images below show three generations of horses.



BLOODSTOCK FOR THE 21ST CENTURY™

EQB



Agents & Consultants

ry



Enter Jeff Seder of EQB, a boutique consulting firm.



• Traditional m

• S

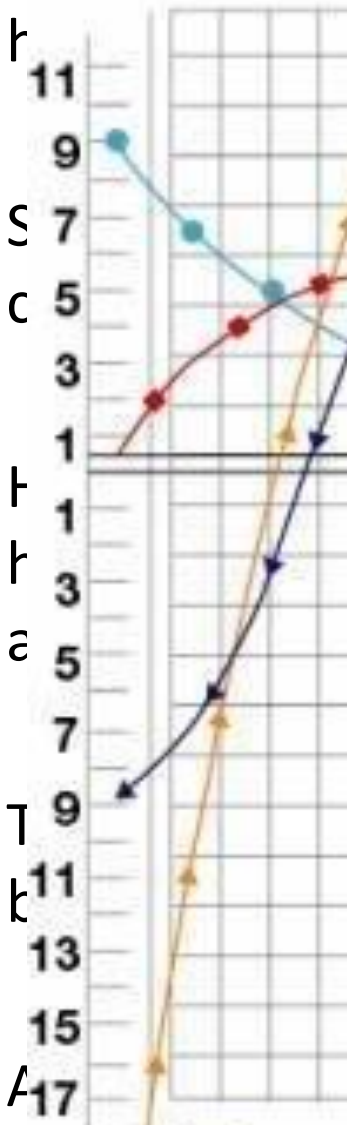
• T

• T

• A

V

VELOCITIES



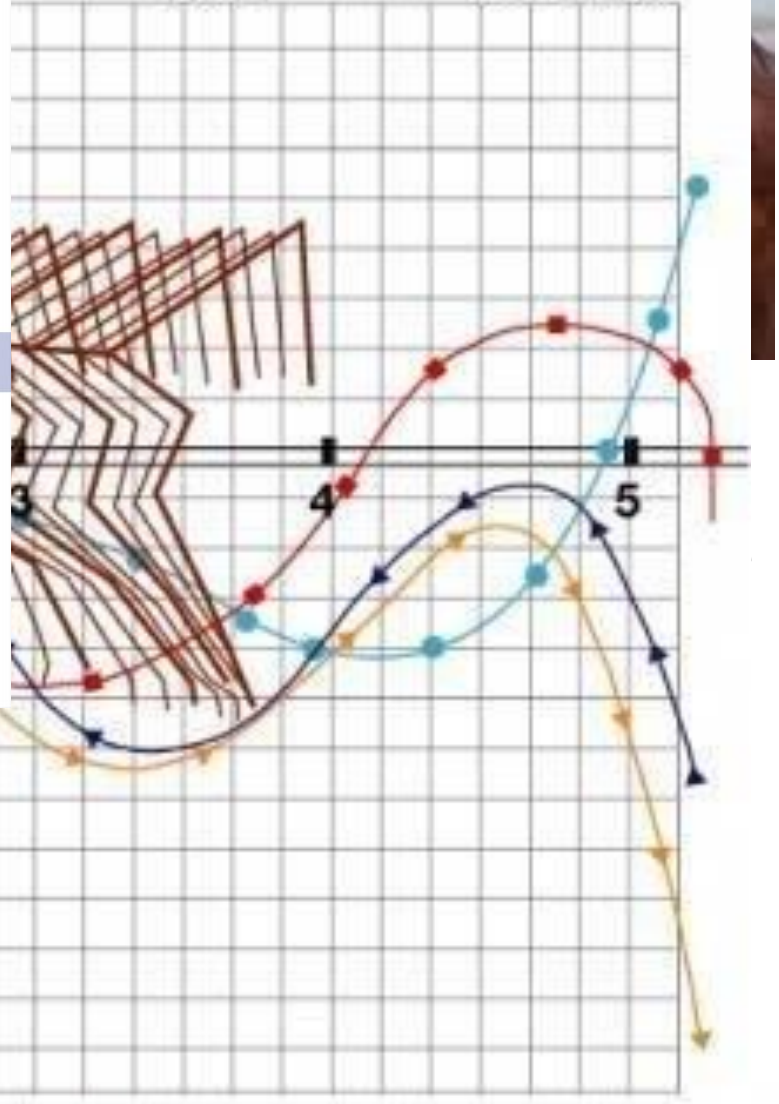
▼ — HOOF ▲ — FETLOCK ■ — LOWER LEG ● — UPPER LEG



ABSOLUTE Medical Equipment



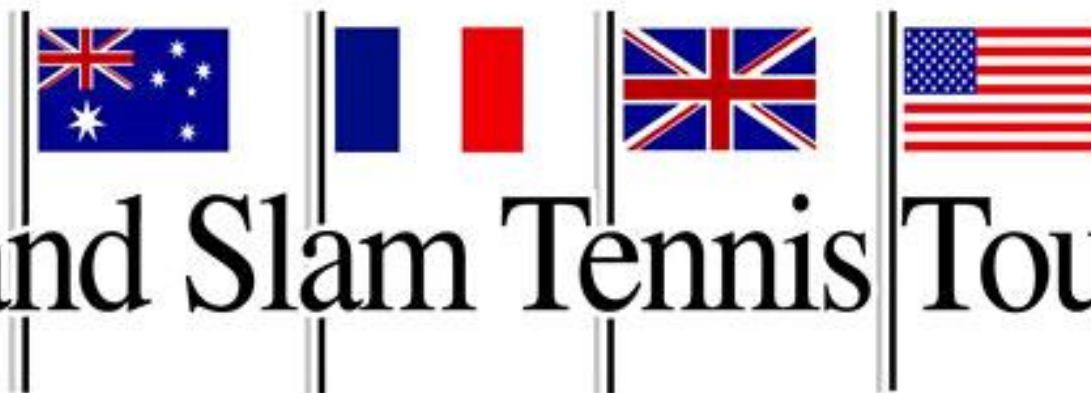
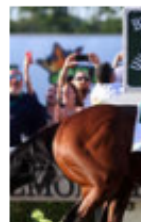
EDDIE SLOWER RACE



A Motivating Example

Am
Cro

JUNE 5, 201



Grand Slam Tennis Tours

day at
in 37

1978. But
ear. OK,

It's hard to bring up Einstein's definition of insanity at this...



raoh became the first horse in 37 years to

A Motivating Example: Concluded

- So, what is the example trying to motivate?
- [1] Importance of having a clear **Objective** to pursue or **Question** to answer.
- [2] **Data is paramount**, when studying, measuring, modeling or understanding any phenomenon of interest.
- [3] Good predictors of an outcome *can* show up in **unexpected places** - where nobody thought to look, **overtaking theories & explanations** - involves **trial-&-error**, guesswork & analytics.
- [4] Important to keep an eye out for **new tech**, which may enable new data to be collected & analyzed.
- [5] Data alone is NOT enough. **Analytics is required**, and an open mindset.

Session Outline

- Motivating Example for Data Analytics
- Preliminaries
- Introduction to Problem Formulation
- Determining Data Requirements
- Some Thoughts on Report Writing: Best Practices
- Session Wrap-up

Some Preliminaries

Preliminaries: About me...

- **Academic Credentials:**

- PhD in Marketing – Univ of Rochester (2009)
- MS in Applied Statistics – Univ of Rochester (2006)
- PGDM – IIM Calcutta (2001)
- B.E. – BIT Mesra (1998)

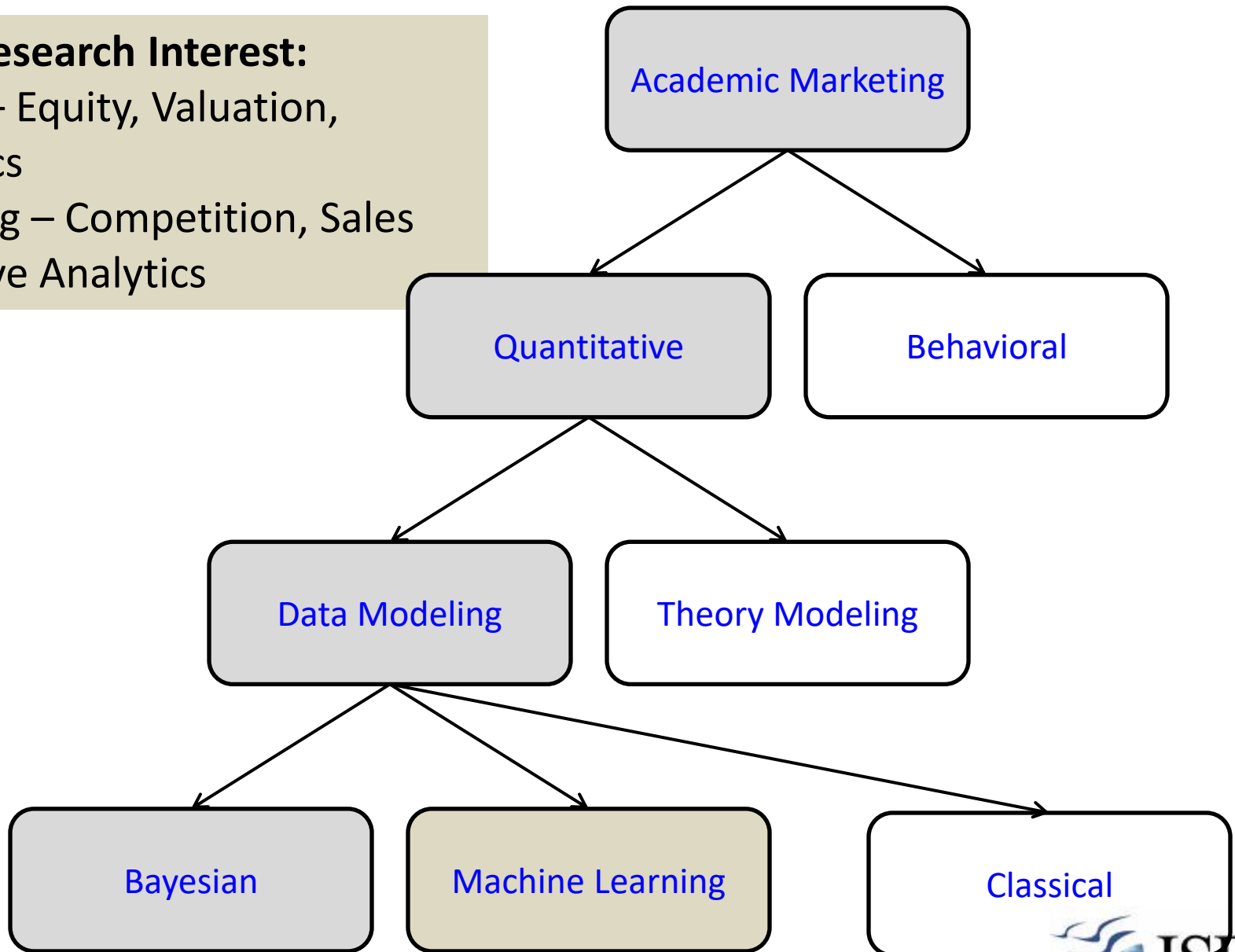
- **Industry Experience:**

- Software Programmer with Cognizant 1998-99
- Management Consultant with Accenture 2001-02
- Data Analyst – Daymon Consumer Insights Division 2006-08
- Academic Faculty with ISB – 2009 onwards
- Been involved in a Tech Startup – Modak Analytics – 2012

Preliminaries: About my Research...

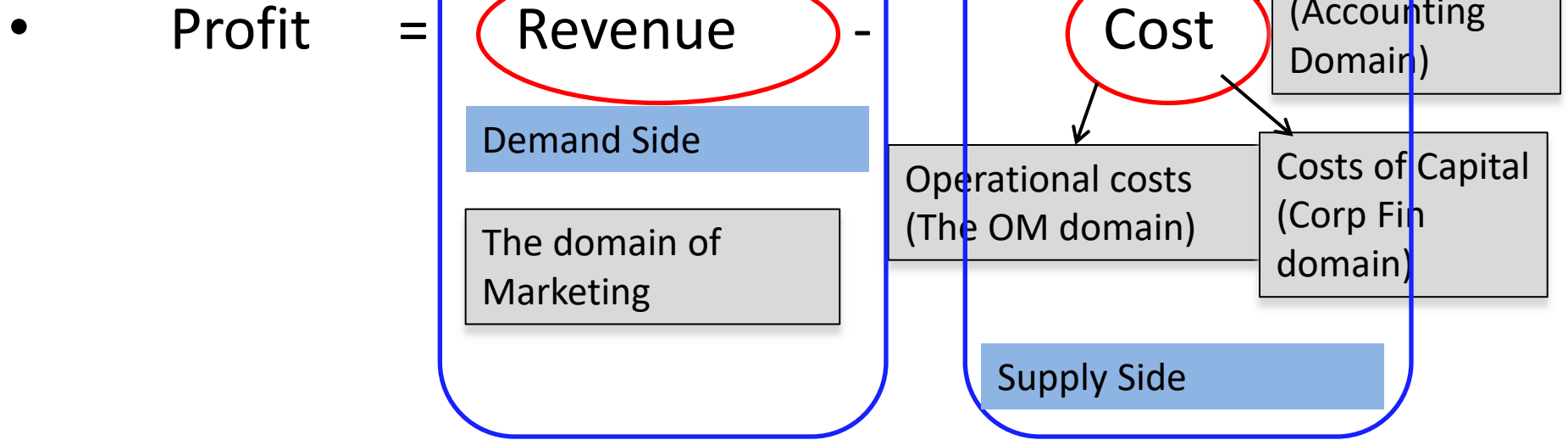
Topics of Research Interest:

1. Brands – Equity, Valuation, Dynamics
2. Modeling – Competition, Sales
3. Predictive Analytics



Preliminaries: The Objective of a Business

- Firms exist to maximize (economic) *profits*



- Business functions represent a logical way to deconstruct the enterprise → yield analytics that is function-specific.
- Market power derives from competencies on either the demand or the supply side.

Preliminaries: The Objectives of Government

- What should government aim for?

- $$\left(\begin{array}{c} \text{Net Societal} \\ \text{Welfare} \end{array} \right) = \left(\begin{array}{c} \text{Consumer} \\ \text{Surplus} \end{array} \right) + \left(\begin{array}{c} \text{Producer} \\ \text{Surplus} \end{array} \right)$$

Ease of citizenry to improve consumption → living standards, at a given price level.

Ease of business to improve production, productivity → profit, at a given price level.

- There is a *tradeoff* between consumer and producer surpluses. If social welfare is constant then raising one means lowering the other.

- Extent of control by government gives us different systems.

Motivating Problem Formulation

Motivating Example

- What's the Mongolian landscape like?



al

SB

Motivating Example

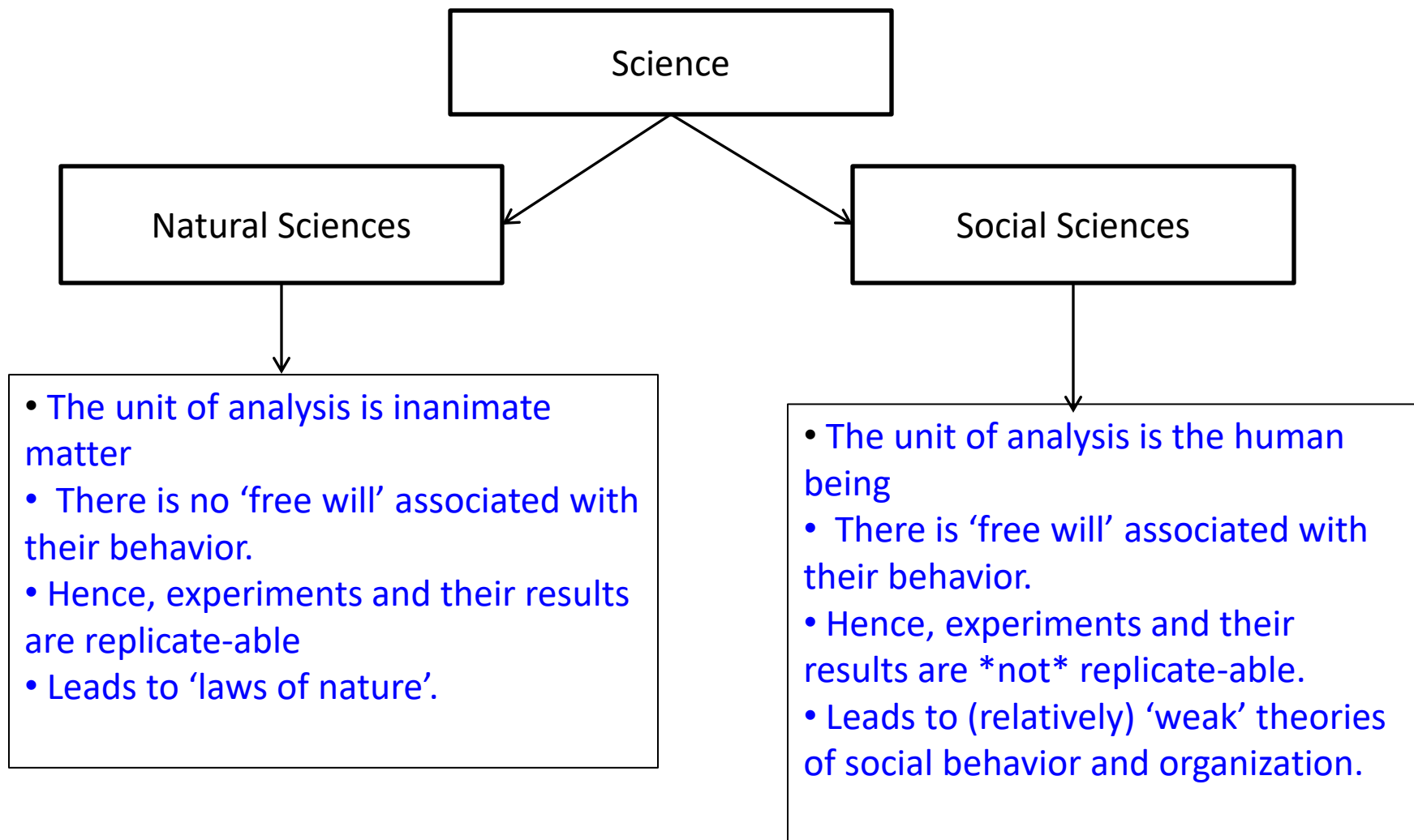
- First, they analyzed the most common diseases needing hospital access.
- Next, they developed DIY (Do-it-Yourself) medicine kits, which like first aid, could be self-medicated after self-diagnosis.
- The DIY kits were placed in each home and their use explained.
- Next, paramedical staff were assigned territories they'd cover once every 6-12 months.
- On each visit, they'd **audit the kit** and the family would **pay only for what medicine was consumed**.
- Simple model, eh? But was it effective? What was the result?

Motivating Example

- Hospital visits declined **45%** in many remote areas → pressure eased on hosp resources and budgets.
- House-call demand for doctors fell **17%** → precious doctor time freed up for other work.
- But more importantly, look at the seemingly simple business model...
- *Medicine as **postpaid** rather than **prepaid**.*
- Extensions? Implications? Further possibilities? Plentiful.
- But remember how it all began... at the problem formulation stage...
- By changing one Q with another, we **transformed** the problem from “**increasing supply of healthcare**” to “**reducing healthcare demand**”...

Conceptual Preliminaries

Preliminaries: Is 'Analytics' Scientific?



Bottomline: There's only so much **precision** in our **measurements** and our results that we can expect. A **Test-and-learn mindset** is critical in the social science side.

Why Identify the Units of Analysis

- Because without **units of analysis**, there is no Measurement.
- Without **Measurement**, there is no Data.
- Without Data, there is no **Analysis**.
- Without Analysis, there is no **Modeling**.
- Without Modeling, there is no **Explanation** and **Prediction**.
- Without Explanation, there is no **Insight**.
- Without Prediction, there can be no **Optimization**.
- Without Insight & Optimization, there is no **Management**.

Test-and-Learn: Evidence based Approach

Home / India / I

How

re

Issu
tout

Writt
Upda



INTERVENTION



The Data Story and History

On Data today

- The **volume, variety and velocity** (the famous three Vs of big data) of the data currently being captured is unprecedented.
- In the time it takes you to read this sentence (~ 6 seconds for the average reader), Google receives **half a million queries** from around the world.
- In 2000, digitally stored data was a mere 25% of all data generated. By 2007, it jumped to 94% (and hasn't fallen since).
- Traditionally, **Data analysis (say, D.A.)** would adapt to whatever data form was available --> D.A. adapted to **D.C. (Data Collection)** --> In turn, D.C. adapted to **Data Generation (say, D.G.)**.
- But the jump from Y2K to 2007 suggests something way more profound.... that **perhaps D.G. is adapting to D.C. is adapting to D.A.?**

Data and the Human Mind

MIND CHANGE



How digital technologies
are leaving their mark
on our brains

SUSAN GREENFIELD

'Britain's best known neuroscientist'
GUARDIAN

te down anything - pen to paper - to

il, go

we k
iline

func

now

on, time horizon perceptions, value
o the web remains to be seen.

Consider the effect of
always-on social network
access, binge-consumption
of video games, audiovisual
entertainment etc. and
technologies to come...

Q is – are they changing
children's brains? Rewiring
circuits, coping and reward
mechanisms? How about
adult brains?

The Age of Data

"If *Land* was the primary raw material of the **agricultural age**,
and *Iron* that of the **industrial age**,
then *Data* is the primary raw material of the **information age**."

Nice quotation. But what's its practical significance?

Consider this Q:

"How many of our present day laws, institutions, societal norms and governance structures actually derive from the agricultural age?"

The Agricultural Age, Data and Governance

Q: How many of our present day laws, institutions, societal norms and governance structures actually derive from the agricultural age?

First Previous Next Last

ಗ್ರಾಮ ಸಮಿತಿ 3
ತಾಲ್ಲೂಕು ಮಂಡಳಿ
ತಾಲ್ಲೂಕು ಗೌಡೇರಿನೂರು
ಹೋಬಳಿ: ನಂದಿಗರೆ
ಗ್ರಾಮ: ಮಲ್ಲೇನಹಳ್ಳಿ
ಪುಟದ ಕ್ರಮ ಸಂಖ್ಯೆ: 17

ರೆಕಾರ್ಡ್ ಆಫ್ ರೈಟ್ಸ್, ಗೇಜ್ ಮತ್ತು ಪಹಣಿ ಪತ್ರಿಕೆ (R.T.C) ಫಾರಂ ನಂ. 02
Print Page No: 1/1 Village Account Form No. 2
Valid from 13/03/2002 To Till Date

1. ವರ್ಗ ಸಂಖ್ಯೆ 10	3. ವೇತನಗಳು ಒಟ್ಟು ಮೈಲಾಜ್ ಪೂಜ್ ಎರಾಜ್ (ಅ) ಪೂಜ್ ಎರಾಜ್ (ಬ) ಉಳಿದವು	ಎಳೆ ಗುಂಪಿ ಅ 1.13.00.00 1.13.00.00	4. ತಂದಾಯ (ಅ) ಭೂ ತಂದಾಯ (ಬ) ಜೋಡಿ (ಕ) ಮೈಸಲು (ಡ) ನೀರಿನ ದರ	ರೂ. ಪೈ 3.00	9. ಕಟ್ಟಡ ಅಥವಾ ಸ್ವಾಧೀನೀಕರಣ ಪದರು ತಂದಾಯ ಪದರು ಮತ್ತು ನಿಲಾಜ್ *ತಳವಾರ ಇನಾಂತ್ -	ಮೈಲಾಜ್ ಎ ಗುಂ 1.13.00.00	ಪಾವತಿ ಸಂ. 114	10. ಕಟ್ಟಡ ಅಥವಾ ಸ್ವಾಧೀನೀಕರಣ ಪದರು	11. ಇತರ ಪಕ್ಕುಗಳು ಮತ್ತು ಪುರಾಣಿ ಪಕ್ಕುಗಳು: ಪಕ್ಕುಗಳು: ಪುರಾಣಿಗಳು:
2. ಒಟ್ಟು				3.00					
5. ಮಲ್ಲೇನಹಳ್ಳಿ ಸಮಿತಿ	7. ಮದಗಲ ಸಂಖ್ಯೆ	8. ವೇತನವಾರ ಪ್ರಕಾರ ನಿಲಾಜೆಯ ಮೈಲಾಜ್							
	ಹೆದರು	ಸಂಖ್ಯೆ	ಕ್ರ. ಸಂ.	ನಿಲಾಜೆ ಮೂಲ	ಮೂಲಾಧಾರ	ಹಿಂಜಾರು	ಬಾಣಿಯು	ಒಟ್ಟು	
	1	1	1	1.13.0.0				1.13.0	



The Industrial Age, Data and Governance

Q: How many of our present day laws, institutions, societal norms and governance structures actually derive from the **Industrial** age?

Types of Joint Stock Company



China Tops Global Startup Growth

The government's business-friendly policies aim to counter an economic slowdown

■ Percentage Increase in New Businesses Since 2010

Q: What Drives [US] Economic Growth?

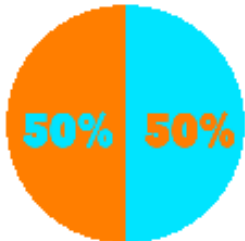
UNITED STATES ECONOMIC ACTIVITY, SPLIT IN HALF

The services sector is the largest (rel. to agri & manufacturing), and much of *growth* in services comes from innovation, from new ideas, materials, methods, technology ...
→ which in turn come from

... Universities. Which require massive funds for both pure and applied research. These funds come from...

... Government. And one of the largest sources for funds within the US govt is the Military.

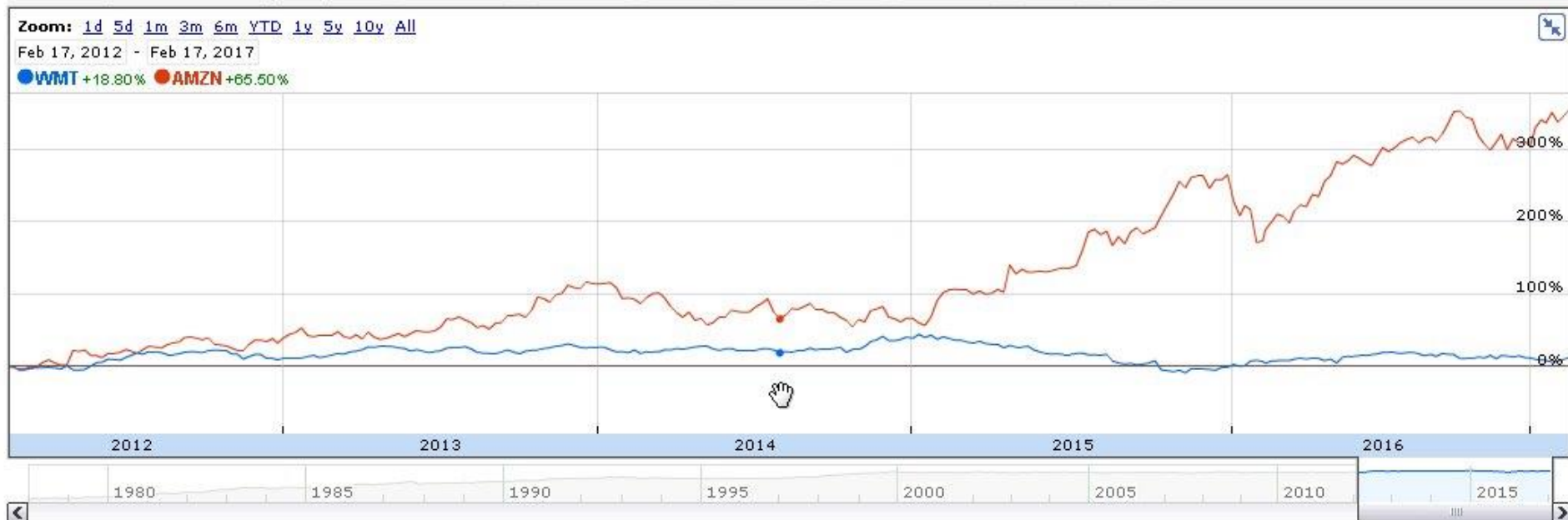
The tiny areas in orange – urban clusters – alone drive 50% of US GDP → Q: What drives economic growth in cities? Consider 3 city clusters...



A. Veltrop

The Information Age, Data and Governance: Example

- Consider the stock performance of Amazon (AMZN) vs Walmart (WMT)



- Valuation, February 2012:
- Walmart: \$202 billion; Amazon: \$82 billion
- Valuation, February 2017:
- Walmart: \$210 billion; Amazon: \$400 billion

Cost of Lost Opportunity: Quick Example

- 2000: Blockbuster had the opportunity to buy Netflix for \$50M

NETFLIX VS. BLOCKBUSTER (2004-2010)



n Disney's

- 2017: @Netflix worth \$61 Bi

Disruption in Action ...

- The world's largest taxi company owns no taxis (Uber)
- The largest accommodation provider owns no rooms (Airbnb)
- Largest phone co.s own no telco infra (Skype, WeChat)
- World's most valuable media firm creates no content (Facebook)
- The world's largest Movie house owns no theatres (Netflix)
- The world's largest software vendors don't write their own code (Apple, Google)
- Etc.

How does Disruption happen?



The next big thing will start out looking like a toy

One of the amazing things about the internet economy is how different the list of top internet properties today looks from **the list ten years ago**. It wasn't as if those former top companies were complacent – most of them acquired and built products like crazy to avoid being displaced.

The reason big new things sneak by incumbents is that **the next big thing always starts out being dismissed as a “toy.”** This is one of the main insights of Clay Christensen's “disruptive technology” theory. This theory starts with the observation that technologies tend to get better at a faster rate than users' needs increase. From this simple insight follows all kinds of interesting conclusions about how markets and products change over time.

Data an

Technology News / Latest Technology News / Startups

Startups » Data localisation payments data digital payment

How local compan



CNBC

Join Now! It's Free



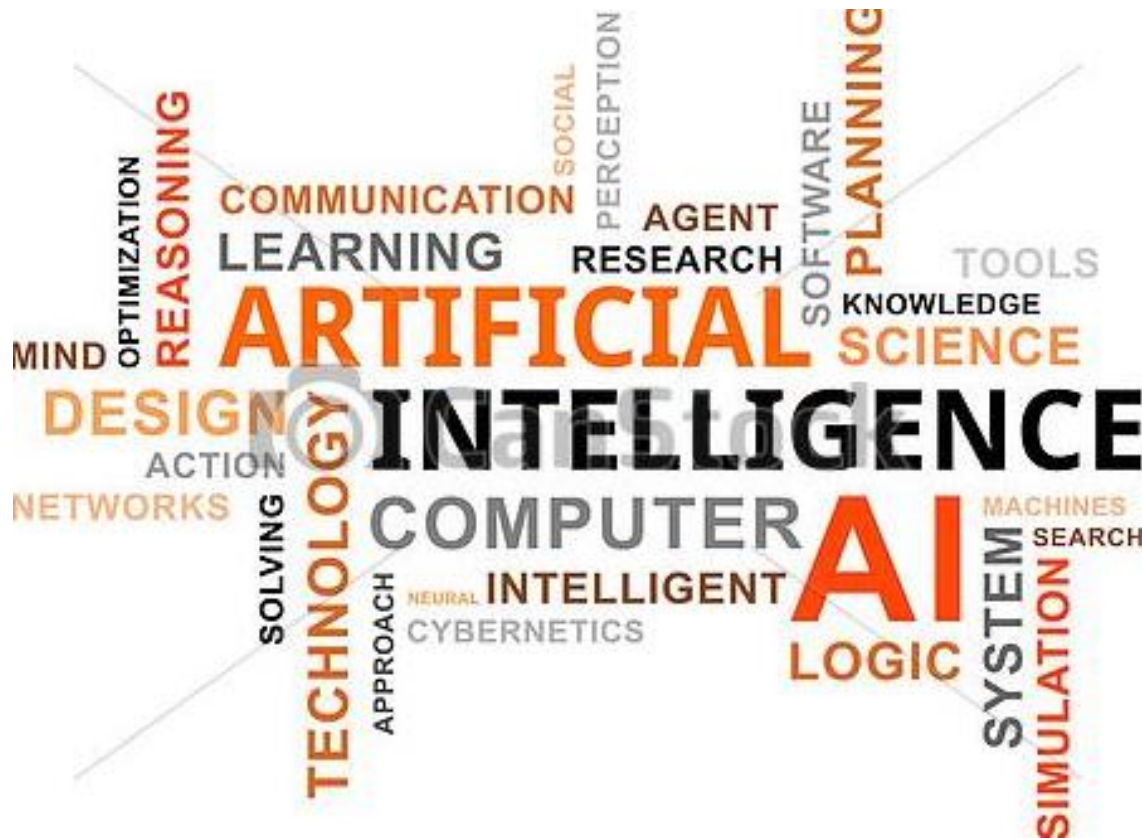
BGR

ISB

Data in the Information Age: The Exponential Learning Curve

The Exponential Learning Curve

- ‘Data Analytics’ often leads to other terms such as ‘machine learning’, ‘artificial intelligence’, ‘blockchain’, etc.
- So what do they mean anyway? How about an example to start figuring out what and how machines *learn* in this century?



The Exponential Learning Curve

- Till 1954, it was widely believed that human beings couldn't run 1 mile in 4 minutes of less. *Why?*
- In 1954, Roger Bannister broke that barrier.

- By 1
- ... W
- bre
- Wh
- driv
- Let



model

ita

The Exponential Learning Curve

- March 13, 2004. The Mojave desert, Calif., site of the DARPA Grand challenge. \$1 million prize money.
- 150 mile race course, numerous [small] obstacles. 15 participants.
- What happened?



The Exponential Learning Curve



Forbes Tech News @ForbesTech - 10m

Toyota says it needs to log at least 8.8 million test miles to ensure autonomous cars are safe on.forbes.com/6018BhsRs



-

- Novem
- This tir
- Cars m
- merge
- What h
- All the
- but wa
- What h
- - Goc
- - Wit
- The po
- Its imp
- cognit
- Implic

The Exponential Learning Curve



A word cloud with 'lifelong-learning' as the largest word. Other prominent words include 'Opportunity independent', 'technology', 'creativity', 'inquiry', 'network', 'fluid fast-paced', 'evolving', 'practise limitless', 'changing', 'future', 'habits', 'countless focus', 'informal-learning', 'self-motivated', 'beyond enthusiasm', 'adapt digital', 'attention', 'extend', and 'independent'.

BRUCE LEE

The next 20 yrs will induce far more changes than the last 20 did.

- We're all destined for lifelong learning, in this lifetime.

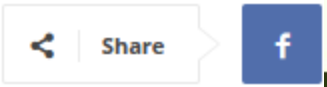
The Cognitive Imperative: Parting take-aways

- We are live-witnessing v interesting times - brimming with profound, radical change.
- Much of what was **common knowledge** in the last century will increasingly **fail** in exploratory, explanatory and predictive power.
- What will replace them will be increasingly **algorithmic**, **impersonal** yet personal, private yet public, **intelligent** yet non-comprehending.
- E.g., Institutions underpin our society, orgs underpin institutions and contractual agreements underpin orgs.
- Think of what happens when 20 yrs on, *blockchains* become as ubiquitous as mobile networks today?
- It is **imperative** we prepare for these sea-changes to come - tool up in terms of **skills, mindset, perspective** ...

Moving from Bits to Atoms ...

- Cons **Indian Govt**

Turtle **policy 2.0**
Open-sourc By **News Desk** - 01/21/2019 3



separate set of Civil Av
under CAR 1.0".

"India is set to become a global leader as far as the drone ecosystem is concerned. It's important for us to have a policy road map and regulations that support the growth of the drone ecosystem," said Jayant Sinha, MoS for Civil Aviation.

- Imp Some specific features of the new drone policy are:

भारत सरकार
GOVERNMENT OF INDIA

रक्षा मंत्रालय
MINISTRY OF DEFENCE



रक्षा उत्पादन विभाग
DEPARTMENT OF
DEFENCE PRODUCTION



Home >>

Defence India Startup Challenge



Welcome to the future.

ms that enable
re, and the
n +

Low-Tech Analytics

The iCow story

- Ve
- Pr
- du
- (a
- (b
- (c
- M
- to
- Si
- Er



National Park

Google

Lamu

Mpeketoni

Map data



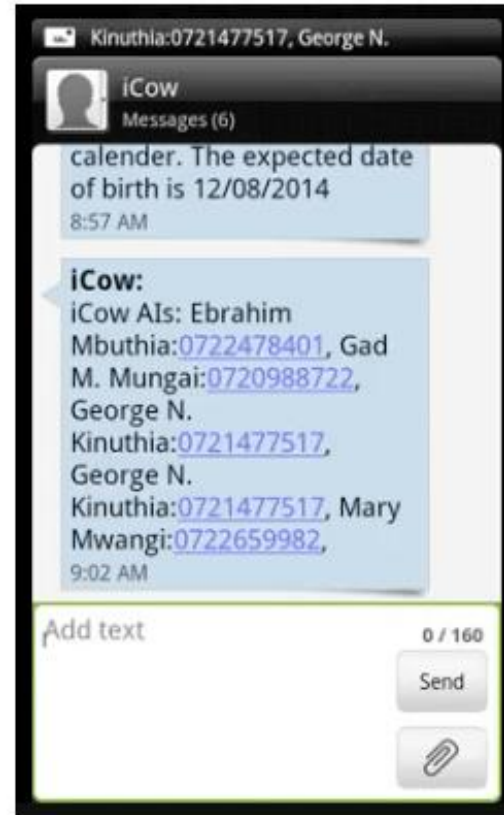
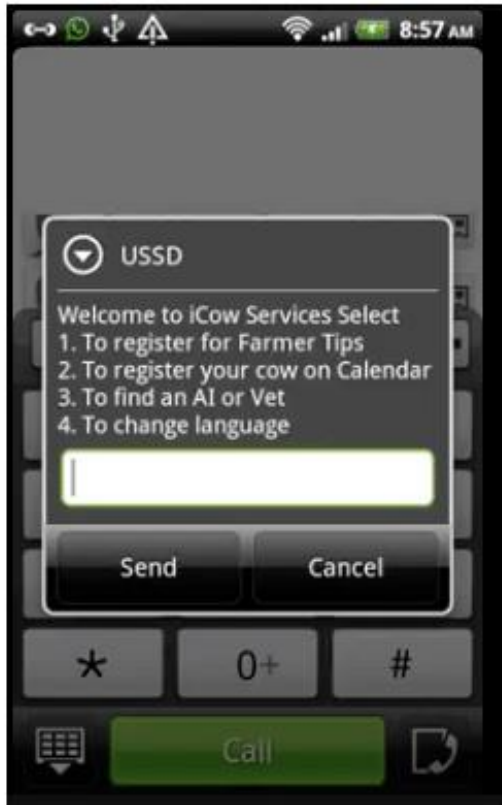
Motivating the iCow story

- iCow says "SMS me info on all 3 issues in standardized format. I'll SMS back instructions to maximize milk yield."

• Word s

• Think c
yield d
nation

• Q: So h
only th

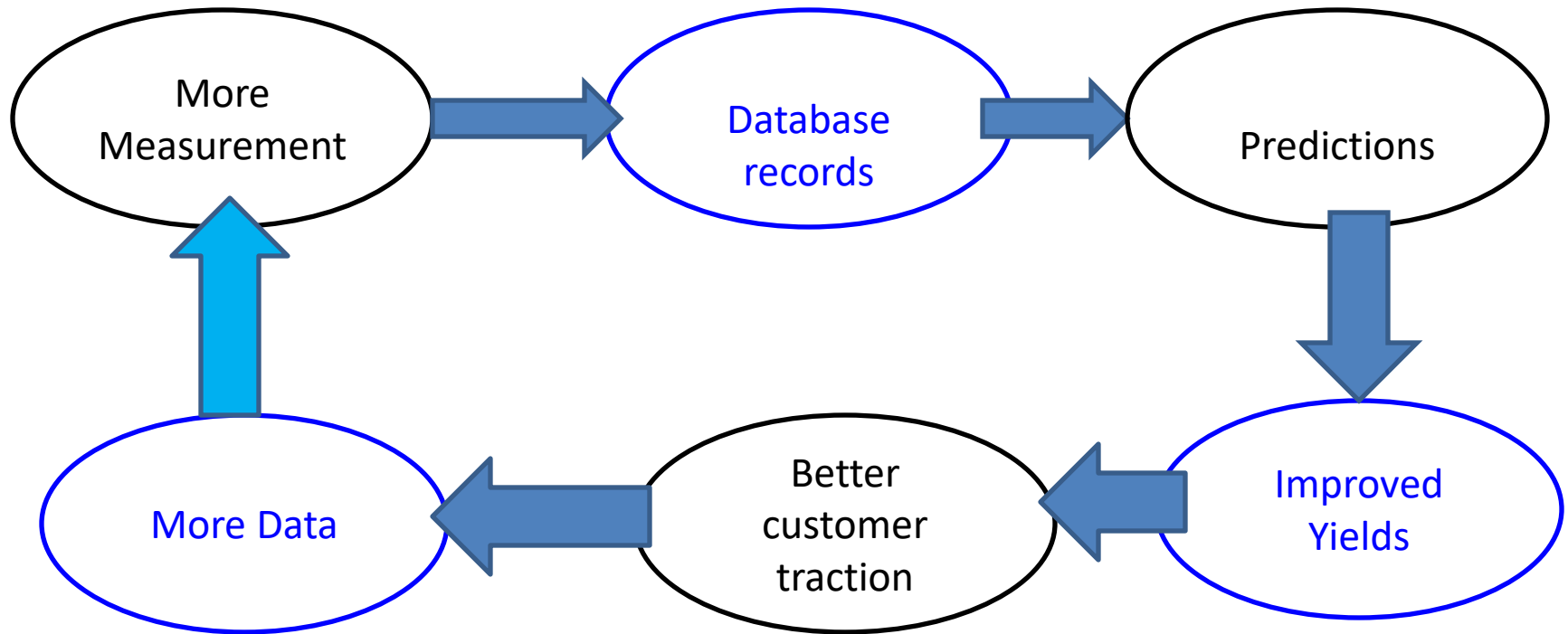


; start tuning in.

do for you &
on
e cow estrus cycle
nd AI Agents
nutrition ng
ing
roduction
related diseases
and Calf care
d diseases

The iCow story: A Virtuous Cycle

- In the beginning, she starts with little or no data and relies primarily on theory and guesswork Later, when the data flow in, analytics is in.



Hence, more signups!

1. Bundled Service

We offer you the ability to bundle iCow with your service or product.

- As i bus We are currently working with micro finance companies that offer financial products to smallholder farmers. I products they are not only adding value to their customer, but also reducing their risk of non performing loans: less of a risk than a farmer without adequate production knowledge.

- (a)
- (b) **2. Marketing**

- (c) We offer marketing opportunities on the iCow platform as well as on the iCow website and social media sites
- (d)

3. Surveys

- So i We offer baseline surveys, longitudinal surveys, analytics, and reporting through our customer care and reser
- (i) i our partners to create in depth knowledge of their beneficiaries.
- (ii)

4. Sponsored content

- Q: \ We offer organisations the ability to add value to their customers with pre-paid iCow products.

to secure Africas food production



What is iCow?

iCow is an Agricultural Information Service with a variety of products available as a subscription service through *285# to help farmers enhance productivity



iCow Products

- Mashauri-Farmer Tips
- Upon subscription to Mashauri-Farmer Tips farmers receive 3 SMS tips per week at Ksh.3 per SMS
- Kalenda-Cow Calendar

[Read More](#)

Basic usage instructions:

Simply Dial *285# and follow the simple menu

iCow Forum



iCow example: Conclusion

- Kahumbu had a very **clear problem statement** in mind:
 - *Optimize* dairy yield (Y) given input variables X.

Problem Formulation!

- To do this, she (ii) needs **data on inputs and yield**, and (ii) needs to **predict What-ifs**.
- To predict What-ifs, she needs a **model for her data**. A model that '*learns*' from data.
- **Data at the individual cow level** - tracing cow's *history* in structured form.
- Individual level data allows for **individual level predictions**.
- Algorithms mine the data for **patterns** and **match** current with patterns for predictions.

Primary Data Collection

We'll build, test and run supervised ML models in MLBM (Term 9)

- The infra and tech required for this operation is now fairly commonplace. And available at **low-cost in open source**.

Thank You

Q & A