1st ASPPH Academic Regional Meeting in Asia:
Global Conference on Public Health Education in the 21st Century.
Organize by:
Association of Schools and Programs of Public Health (ASPPH) and
The National Taiwan University College of Public Health (NTU).

Title:

Big Data Science: Does It in Public Health.

By:

Datuk Dr Mohd Yusof Ibrahim.
MD; MPH (MCH);AM(Mal).
Senior Consultant Public Health Physician,
Head of Community and family Medicine,
Faculty of Medicine and Health Sciences,
Universiti Malaysia Sabah (UMS).
Malaysia.

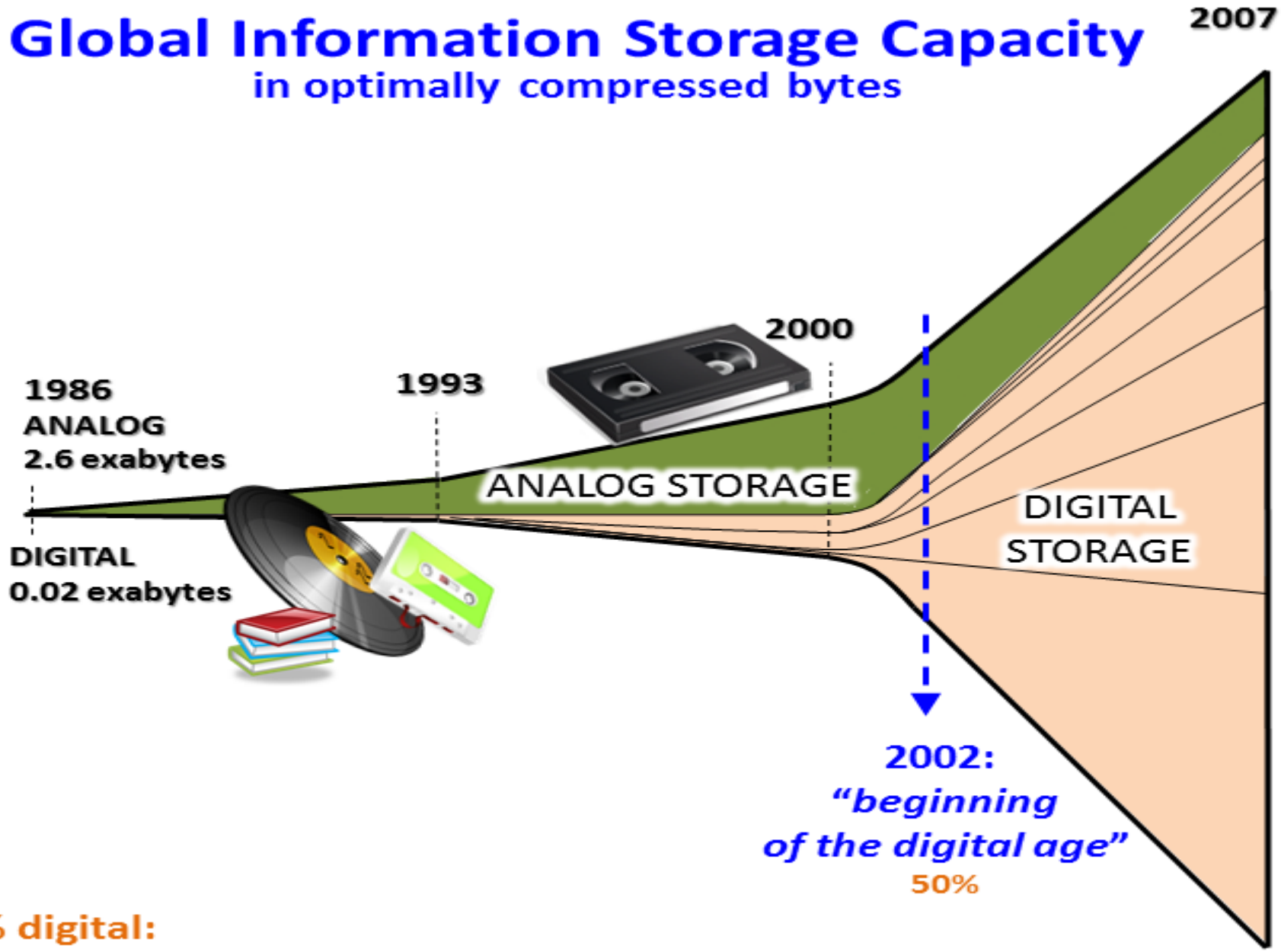
24 September 2019.

Outline Presentation

- WHAT IS BIG DATA
- APPLICATIONS
- CHALLANGES
- FUTURE DIRECTION
- CONCLUSION

Global Information Storage Capacity

in optimally compressed bytes



2007

ANALOG
19 exabytes

- Paper, film, audiotape and vinyl: 6 %
- Analog videotapes (VHS, etc): 94 %

ANALOG

DIGITAL

- Portable media, flash drives: 2 %
- Portable hard disks: 2.4 %
- CDs and minidisks: 6.8 %
- Computer servers and mainframes: 8.9 %
- Digital tape: 11.8 %
- DVD/Blu-ray: 22.8 %

- PC hard disks: 44.5 %
  123 billion gigabytes

- Others: < 1 % (incl. chip cards, memory cards, floppy disks, mobile phones, PDAs, cameras/camcorders, video games)

DIGITAL
280 exabytes

Source: Hilbert, M., & López, P. (2011). The World's Technological Capacity to Store, Communicate, and Compute Information. *Science*, 332(6025), 60 –65. http://www.martinhilbert.net/WorldInfoCapacity.html

Terms

- e-health,
- m-health,
- digital health,
- health information technology,
- health 2.0,
- e-medicine,
- telemedicine,
- e-health

Big Data in Health

- *"refers to large routinely or automatically **collected datasets**, which are **electronically captured** and **stored**.*
- *It is **reusable** in the sense of multipurpose data and comprises the fusion and connection of existing databases **for the purpose of improving health and health system performance.***
- *It does not refer to data collected for a specific study."*

Big data ?

- means different things to different people.
- in a health context **big data**=
  - **large databases** where our interactions with the health care system are stored.
- **5 descriptions:**
  1. measures of **participant biology**, as in genomic or metabolomic data sets;
  2. measures of **participant context**, as in geospatial analyses
  3. administratively collected **medical & Health record data**
  4. participant **measurements taken automatically at extremely frequent intervals as by a global positioning system** (GPS) device
  5. measures **compiled from the data effluent** created by life in an electronic world, such as search term records

Characteristics Of Datasets To Be Considered As Big Data

Velocity

- Data speed

Volume

- **Data quantity**

Terabytes
Zettabytes
Brontobytes.

Variety

- Data Types

VERACITY

messiness

Turning Big Data into Value

The 'Datafication' of our World;

- Activities
- Conversations
- Words
- Voice
- Social Media
- Browser logs
- Photos
- Videos
- Sensors
- Etc.

Volume

Velocity

Variety

Veracity

Analysing Big Data:

- Text analytics
- Sentiment analysis
- Face recognition
- Voice analytics
- Movement analytics
- Etc.

Sources of Big Data in Health :

- Health care records and patient summaries
- Social media
- Genomic data
- Pharmaceutical data
- Insurance claims
- Telemedicine, mobile apps and sensors
- Other sources (income statistics, environmental data bases etc.)

Big Data in Healthcare: Why?

- **Big Data enables to test exhaustively these claims:**
- ❑ Faster identification of high-risk patients
- ❑ More effective interventions
- ❑ Better decision making
- ❑ Closer monitoring
- ❑ Correlate “prior” clinical knowledge vs Big Data crunching

Big Data in Healthcare: How?

- ❑ Massive storage
  - Increasingly detailed data for each individual—including genomic,
  - cellular, environmental data, historical patient records, clinical trials
- ❑ Massive computation
  - Distributed and parallel computation on commodity hardware
- ❑ Powerful analytics
  - Allowing to process large amounts of data in batch or in streaming

Big Data in Healthcare: Benefits

Right living

Informed lifestyle choices that promote well-being and the active engagement of consumers in their own care

Right provider

Care provider (eg, nurse, physician) and setting that is most appropriate to deliver prescribed clinical impact

Right care

Evidence-based care that is proven to deliver needed outcomes for each patient while ensuring safety

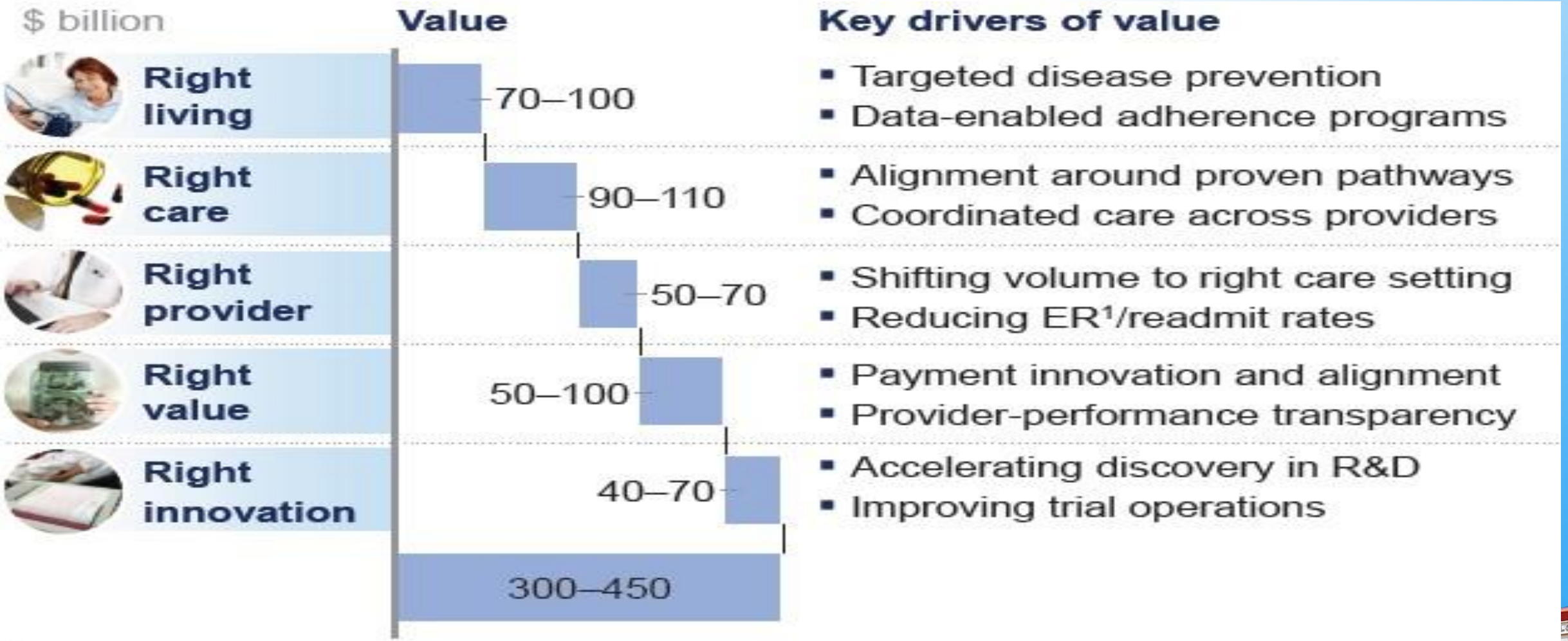
Right value

Sustainable approaches that continuously enhance healthcare value by reducing cost at the same or better quality

Right innovation

Innovation to advance the frontiers of medicine and boost R&D productivity in discovery, development, and safety

Big Data in Healthcare:
Economic benefits



1 Emergency room.

Source: American Diabetes Association; American Hospital Association; HealthPartners Research Foundation; McKinsey Global Institute; National Bureau of Economic Research; US Census Bureau

Applications

1. increasing the **effectiveness and quality of treatments**
2. **prevention of diseases**
    1. **identification of risk factors for disease at population, subpopulation, and individual levels, and by improving the effectiveness of interventions to**
3. **prediction of outcomes /**predictive models.
    1. **containment and improvement of chronic diseases, global infectious disease surveillance**
    2. **better understanding of demographic challenges**
    3. **trends / transmission pathways.**
4. **knowledge dissemination**
    1. **to stay current with the latest evidence guiding clinical practice.**
5. **reduction in inefficiency and waste.**
6. **improvement in cost containment.**

BIG DATA SURVEILLANCE

- Public health surveillance systems
  - monitor **trends in disease incidence**, health behaviours and environmental conditions in order to allocate resources to maintain healthy populations
- **prediction** or hypothesis generation
  - rather than hypothesis testing .

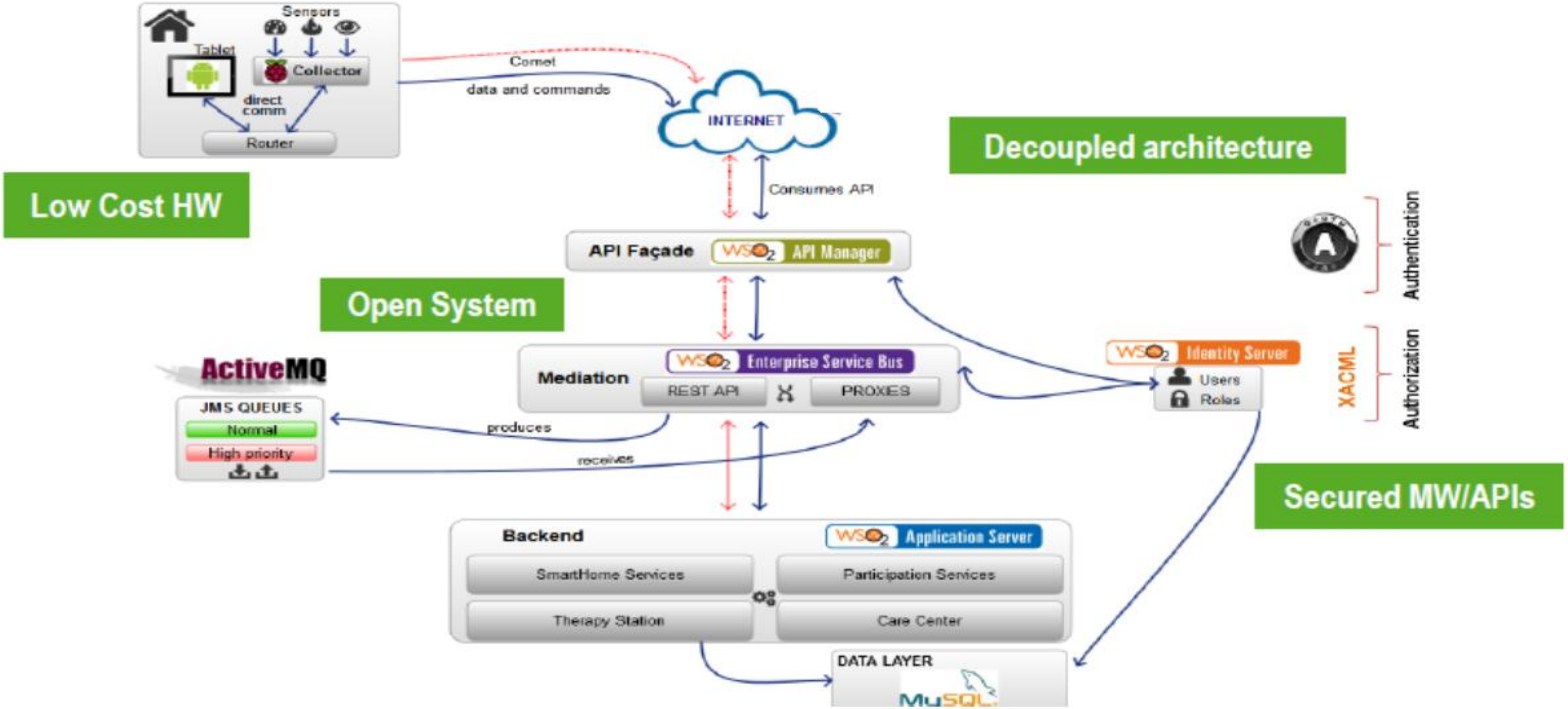
Other uses:….

- **outbreak detection** and surveillance
- identification of **patient features** associated with clinical outcomes
- **environmental monitoring**
- build an **early warning system** for adverse drug reactions from social media data
- to **detect falls** from Smartphone data
- to **identify outlier** air pollutants
- to **predict hospital readmission**
- tuberculosis **transmission**
- **serious injuries** in motor vehicle crashes
- **suicidal ideation**
- .

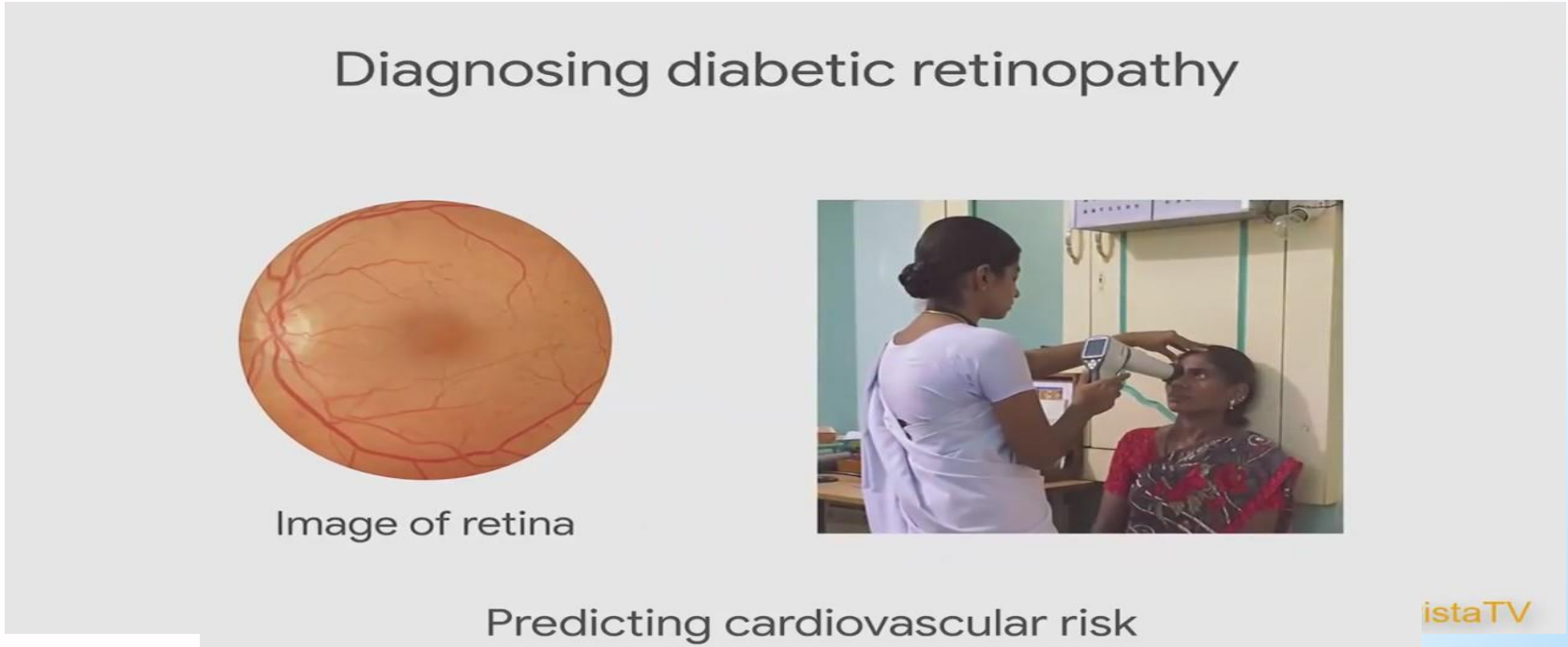
Application: Case 1- Advanced Teleassistance System

- People **aged from 65-80 will raise by nearly 30% in 2020**
- How to improve **elderly quality of life, autonomy and independence given heath system** limitations
  - Required **improve teleassistance services to prevent emergency situations, early detection and** prevention of aging Illnesses (activity decline, alzheimer, dementia)
- **Variety of data:**
  - Home automation sensor data (e.g. door switches, motion sensors, environnemental sensors, gas, smoke detectors)
- **Real time Data Analysis Challenges:**
  - Activity/Inactivity
  - Provide summary of action events (sleep hours)
  - Fit models with variety of data

Use case: Advanced teleassistance system



Application: Case 2-
Community Diagnosing Diabetic Retinopathy



Application: Case 3- Predicating Medical Events

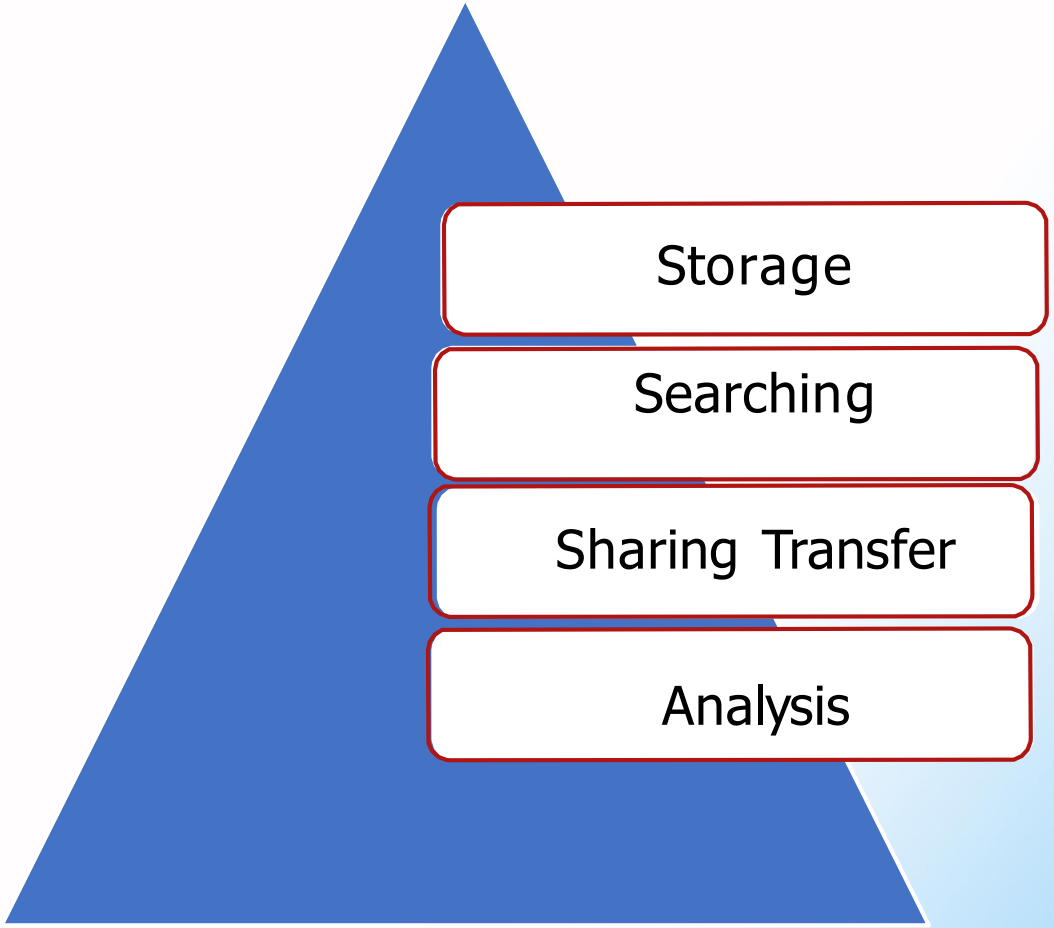
In Malaysia…

❑DG MOH – New technology : AI statescope

❑Predicting Dengue outbreak

- ❑ Predicting Malaria outbreak
- ❑ Predicting success of TB Treatment
- ❑ Predicting of AGE
- ❑ etc……

Challenges


Storage
Searching
Sharing Transfer
Analysis


DEWAN CANSELOR

Challengers

- **Security & privacy.**
  - **Concentrate and analyse so much data gives** too much **power and responsibility to the institutions,** organisations, private entities in terms of Information that can be derived or leaked
- **Data analysis challenge.**
  - **Data is the new oil, but massive data** sometimes difficult and messy to extract
- **Cost.**
  - **One of the biggest stumbling blocks for Big Data in** healthcare right now, even though applications like Hadoop are open-source
- **Recent technologies.**
  - **Continuous tools in market, new tools** require consolidation, few experts and difficult to find

Recommendations

1. **Awareness Raising:**
   - ❑ Develop and implement a communication strategy to increase the awareness of the added value of Big Data in Health and encourage a positive public mind set towards Big Data in Health.
2. **Education and Training:**
   - ❑ Strengthen human capital with respect to the increasing need for a workforce that can utilize the potential of Big Data in Health
3. **Data Sources:**
   - ❑ Expand existing and explore new sources of Big Data in Health and secure their quality and safety
4. **Open Data and Data Sharing:**
   - ❑ Promote open use and sharing of Big Data without compromising patients' rights to privacy and confidentiality
5. **Applications and Purposes:**
   - ❑ Target-oriented application of Big Data analysis in health based on the needs and interests of stakeholders including patients

Cont...

6. **Data Analysis:**
   - ❑ Identify the potentials of Big Data analysis, improve analytical methods and facilitate the use of new and innovative analytical methods.
7. **Governance of Data Access and Use:**
   - ❑ Implement governance mechanisms to ensure secure and fair access and use of Big Data for research in health.
8. **Standards:**
   - ❑ Develop standards for Big Data in Health to enhance and simplify its application and improve interoperability
9. **Funding and Financial Resources:**
   - ❑ Ensure purposeful investment steered by the European Commission to war-rant cost-effectiveness and sustainability
10. **Legal Aspects and Privacy Regulations:**
    - ❑ Clarify and align existing legal and privacy regulation of Big Data in Health.

LIMITATIONS AND ISSUES

1. machine learning's capacity to overcome the curse of dimensionality **requires tall data sets**
2. machine-learning models are often described as "black boxes" whose opacity **precludes interpretability** or sanity-checking of key assumptions by nonexperts
3. observers **assume** that models that learn automatically from **data are more objective** and therefore **more accurate** than human-constructed models.

FUTURE DIRECTIONS

1. Develop the capacity to **think like a computer** when working with data
2. Public health training, especially within epidemiology and biostatistics:
   - ❑ Esp. quantitative bias analysis and related techniques
   - ❑ including an understanding of health informatics, data engineering, computational complexity, and adaptive learning.

Conclusions

- Big data provides infrastructure and analytics for massive data
- - Big Data can improve healthcare system
- - Research can boost its productivity
- - Key technical aspects- Big Data raises security and privacy concerns
- - Required right policies to make the change possible in healthcare.

Applications Of Big Data Are Endless!

Any business that doesn't seriously consider the implications of Big Data ….

runs the risk of being left behind.


Star TV
NEWS


DEWAN CANSELOR

REF

1. Barrett MA, Humblet O, Hiatt RA, Adler NE. Big Data and Disease Prevention: From Quantified Self to Quantified Communities. Big Data. 2013;1(3): 168-75Stephen J. Mooney[1] and Vikas Pejaver[2]. Annual Review of Public Health Vol. 39:95-112 (Volume publication date April 2018)
2. Bernard Marr, John Wiley & Sons. Big Data in Practice. 2 May 2016
3. Edwin Morley-Fletcher. An overview of the challenges in data intensive healthcare
4. Gesundheit Österreich Forschungs- und Planungs GmbH *December.* Study on Big Data in Public Health, Telemedcine and Healthcare. Final Report . December 2016 EUROPEAN COMMISSION .Directorate-General for Health and Food Safety. Directorate B — Health systems, medical products and innovation *European Commission B-1049 Brussels, 2016*
5. IBM. The four V's of Big Data.: Hub IBDaA, editor. Info graphics & Animations: http://www.ibmbigdatahub.com/infographic/four-vs-big-data.
6. Jared Dean. Big Data, data mining and Machine Learning. Wiley, 2014
7. Jure Leskovec, Anand Rajaraman, Jeffrey D. Ullman. Mining of Massive Datasets. 2014
8. Muin J. Khoury and John P. A. Ioannidis. Big data meets public health. Science. 2014 November 28; 346(6213): 1054–1055. doi:10.1126/science.2709
9. Peter Groves, Basel Kayyali, David Knott, Steve Van Kuiken. The ‘big data’ revolution in healthcare. Accelerating value and innovation. Centre for US Health System Reform. Business Technology Office. 2013 Frontiers in Massive Data Analysis. THE NATIONAL ACADEMIES PRESS, 2013
10. Vignesh Prajapati. Big Data Analytics with R and Hadoop. Packt Publishing, 2013
11. Wang W, Krishnan E. Big Data and Clinicians: A Review on the State of the Science. JMIR Med Inform 2014;2(1):e1
12. WHO. FROM INNOVATION TO IMPLEMENTATION: eHealth in the WHO European Region. Copenhagen: WHO Regional Office for Europe, 2016