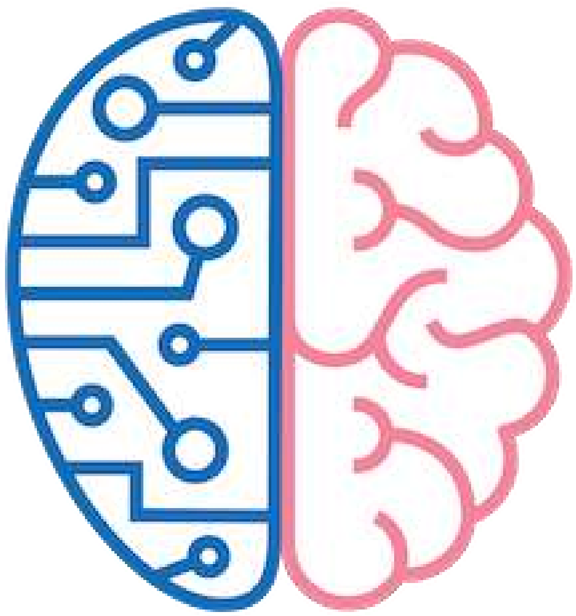


Medical Intelligence

Principles and Applications of Artificial Intelligence in Medicine and Healthcare

Anthony Chang, MD, MBA, MPH, MS
Chief Intelligence and Innovation Officer
Medical Director, The Sharon Disney Lund
Medical Intelligence and Innovation Institute (MI3)
Children's Hospital of Orange County
Computer Scientist in Residence
Chapman University

Anthony C. Chang, MD, MBA, MPH, MS

Dr. Chang attended Johns Hopkins University for his B.A. in molecular biology prior to entering Georgetown University School of Medicine for his M.D. He then completed his pediatric residency at Children's Hospital National Medical Center and his pediatric cardiology fellowship at the Children's Hospital of Philadelphia. He then accepted a position as attending cardiologist in the cardiovascular intensive care unit of Boston Children's Hospital and as assistant professor at Harvard Medical School. He has been the medical director of several pediatric cardiac intensive care programs (including Children's Hospital of Los Angeles, Miami Children's Hospital, and Texas Children's Hospital). He served as the medical director of the Heart Institute at Children's Hospital of Orange County.

He is currently the Chief Intelligence and Innovation Officer (CIIO) and Medical Director of the Heart Failure Program at Children's Hospital of Orange County. He has also been named a Physician of Excellence by the Orange County Medical Association and Top Cardiologist, Top Doctor for many years as well as one of the nation's Top Innovators in Healthcare.

He has completed a Masters in Business Administration (MBA) in Health Care Administration at the University of Miami School of Business and graduated with the McCaw Award of Academic Excellence. He also completed a Masters in Public Health (MPH) in Health Care Policy at the Jonathan Fielding School of Public Health of the University of California, Los Angeles and graduated with the Dean's Award for Academic Excellence. Finally, he graduated with his Masters of Science (MS) in Biomedical Data Science with a subarea focus in artificial intelligence from Stanford School of Medicine and has completed a certification on artificial Intelligence from MIT. He is a computer scientist-in-residence and a member of the Dean's Scientific Council at Chapman University.

He has helped to build a successful cardiology practice as a startup company and was able to complete a deal on Wall Street. He is known for several innovations in pediatric cardiac care, including introducing the cardiac drug milrinone and co-designing (with Dr. Michael DeBakey) an axial-type ventricular assist device in children. He is a committee member of the National Institute of Health pediatric grant review committee. He is the editor of several textbooks in pediatric cardiology, including *Pediatric Cardiac Intensive Care, Heart Failure in Children and Young Adults*, and *Pediatric Cardiology Board Review*.

He is the founder of the Pediatric Cardiac Intensive Care Society (PCICS) that launched the multi-disciplinary focus on cardiac intensive care for children. He is also the founder of the Asia-Pacific Pediatric Cardiac Society (APPCS), which united pediatric cardiologists and cardiac surgeons from 24 Asian countries and launched a biennial meeting in Asia that now draws over 1,000 attendees.

He is the founder and medical director of the Medical Intelligence and Innovation Institute (MI3) that is supported by the Sharon Disney Lund Foundation. The institute is dedicated to implement data science and artificial intelligence in medicine and is the first institute of its kind in a hospital. The new institute is concomitantly dedicated to facilitate innovation in children and health care all over the world. He is the organizing chair for the biennial *Pediatrics2040: Emerging Trends and Future Innovations* meeting as well as the founder and director of the Medical Intelligence and Innovation Summer Internship Program, which mentors close to 100 young physicians-to-be every summer. He has organized a pediatric innovation leadership group called the international Society for Pediatric Innovation (iSPI).

He intends to build a clinician-computer scientist interface to enhance all aspects of data science and artificial intelligence in health and medicine. He currently lectures widely on big data and artificial intelligence in medicine (he has been called "Dr. A.I." by the *Chicago Tribune* and has given a TEDx talk as well as on the Singularity University faculty) ([1]). He has published review papers on big data and predictive analytics as well as machine learning and artificial intelligence in medicine ([2])([3]). He is on the editorial board of the *Journal of Medical Artificial Intelligence.* He is currently completing a book project with a book series with Elsevier: *Medical Intelligence: Principles and Applications of Artificial Intelligence in Medicine and Healthcare.* He is the founder and organizing chair of several *Artificial Intelligence in Medicine (AI Med)* meetings in the U.S. and abroad (Europe and Asia) that will focus on artificial intelligence in healthcare and medicine. He intends to start a new group for clinicians with a special focus on data science and artificial intelligence (MD4ai) as part of a new society (Medical Intelligence Society, or MIS).

He is the CEO and co-founder of three startup companies in artificial intelligence in medicine:
1) CardioGenomic Intelligence (CGI), LLC is a multifaceted company that focuses on artificial intelligence applications such as deep learning in clinical cardiology (cardiomyopathy and heart failure as well as other cardiovascular disease) and genomic medicine.
2) Artificial Intelligence in Medicine (AIMed), LLC is a multi-media company that organizes meetings and educational programs in artificial intelligence in medicine in local as well as global venues.
3) Medical Intelligence (MI7), LLC is an education and consulting/advising conglomerate for executives and physician leaders as well as investors for the evaluation and implementation of AI strategies in healthcare organizations, for evaluation and recommendation of AI in healthcare vendors, and assessment and implementation of cybersecurity in healthcare organizations.

---

[1] https://www.youtube.com/watch?v=Y5T8nckyuCA
[2] www.congenitalcardiologytoday.com/index_files/CCT-NOV12-NA.pdf
[3] www.congenitalcardiologytoday.com/index_files/CCT-APR13-NA.pdf

DEDICATION

The book is dedicated to my daughters Emma and Olivia and the many thousands of children and adults with diseases whom I have had the pleasure to serve as their cardiologist (including a very special nine year-old girl named Elsa from Myanmar) as well as the millions of patients and families and clinicians worldwide who are eternally dedicated to improve healthcare. Their supreme fortitude in their long and complicated medical journeys will continually inspire me to maintain my ardent passion to learn and teach in this nascent and wondrous world of artificial Intelligence in medicine and healthcare.

ACKNOWLEDGEMENT

I would like to express my gratitude to the Sharon Disney Lund Foundation board members, in particular Michelle Lund and Robert and Gloria Wilson. I like to express my thanks also to my Medical Intelligence and Innovation Institute (MI3) team members: Dr. Spyro Mousses, Ms. Laura Beken, Ms. Seraya Martinez, Ms. Vanessa Avina, Ms. Mijanou Pham, Dr. Sharief Taraman, and our advanced fellows in Artificial Intelligence in Medicine, Nathaniel Bischoff and Alex Barrett. I like to thank Ms. Kimberly Cripe, our CEO and Dr. Nick Anas, our pediatrician in chief for their ardent support as well as Dr. William Feaster and Mr. John Henderson along with other members of the IT group, all exemplary in their professionalism.

I would like to thank my Stanford School of Medicine Biomedical Data Science program mentors and classmates for their utmost patience and gracious encouragement. I would also like to thank my computer science colleagues at Chapman University Department of Computer Science, especially Dr. Michael Fahy. I like to express my gratitude to the tireless and dedicated staff of the Artificial Intelligence in Medicine (AIMed) project: Freddy White, Andrew Davies, and Charlie Maloney and the many colleagues and friends who have participated as faculty members as well as attendees at the Artificial Intelligence in Medicine (AIMed) annual meetings around the world. Finally, I would like to thank all those authors who have kindly contributed to this book.

PREFACE

This book with the accompanying compendium is designed for anyone who is interested in a comprehensive primer on the principles and application of artificial intelligence in healthcare and medicine. The readers include the dedicated but busy clinician, the interested data/computer scientist, the astute investor, the inquisitive hospital administrator and leader (from CEO to CIO and others), and any curious patient and family member. In short, it is written with everyone in mind.

The sections of the book are designed to provide a comprehensive framework for anyone who is interested in understanding both artificial intelligence as well as the aspects of artificial intelligence that would be relevant to biomedicine and healthcare. There are commentaries by experts to add a personal dimension to the book

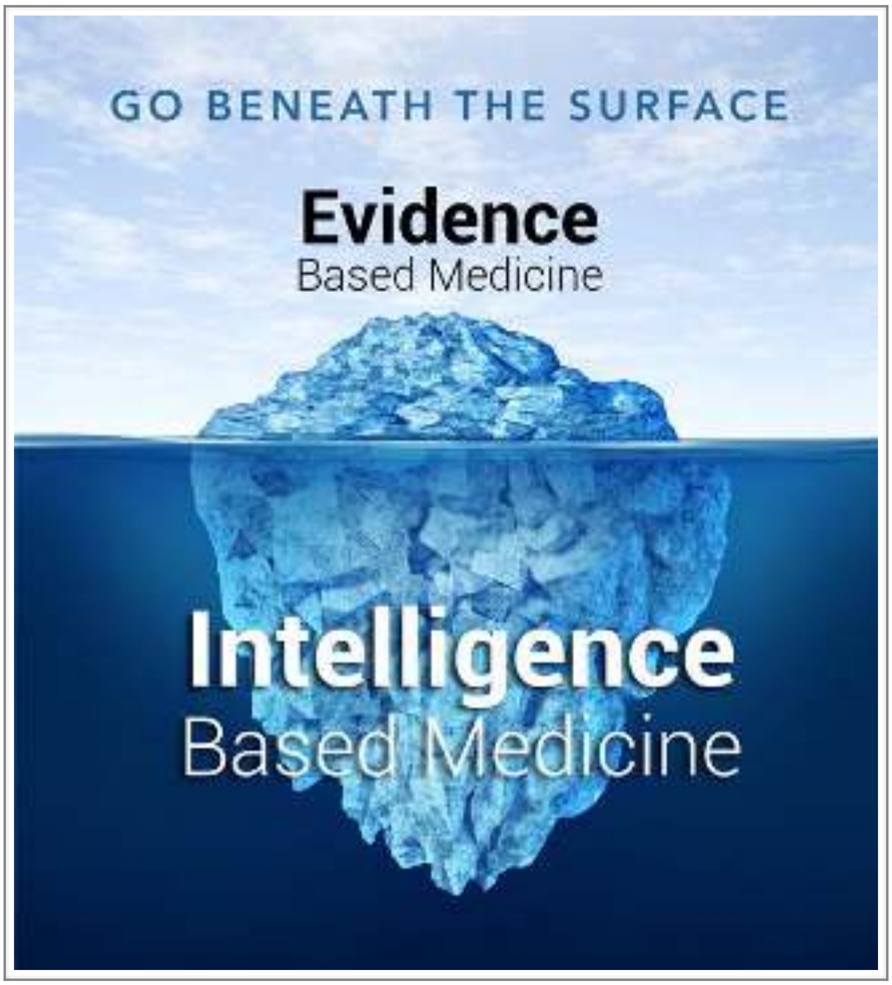
The first Introduction section elucidates the basic concepts of artificial intelligence and explores the relationship between neuroscience and artificial intelligence.

The second section of the book delineates the early years of artificial intelligence and the history of artificial intelligence in medicine during this early era. The genesis of AI and its application is medicine is key to the understanding of state of the art today.

The third section details the current era or state-of-the-art of artificial intelligence as it stands today in 2018 and covers basic elements of algorithms, big data, cognitive computing, and deep learning (the A,B,C,D of the current AI era). Additional key concepts such as natural language processing and internet of things are also included in this section. Following this background of AI, the concepts of artificial intelligence in medicine such as how doctors and machines think as well as the nuances of healthcare data and databases are described. Deep learning, the workhorse of AI in this era, as it relates to healthcare, is separately explained.
Applications of AI in medicine in six areas are then summarized in the subsequent subsections. Following this orientation, AI in medicine as it relates to various subspecialties is then separately discussed for a number of subspecialties.

The future of AI and applications in medicine is then covered in the fourth section of the book. The future of AI is discussed in terms of future elements such as augmented and virtual reality and internet of everything. Additional key concepts such as virtual assistants and quantum computing are also covered. The future of AI as it relates to medicine is separately discussed, followed by a special summary of the ten essential problems and solutions for the AI paradigm in medicine to be successful.

The last section of the book is a compendium of useful resources including: key references, glossary of key terms, tutorial of computer programs, and a list of companies in AI in healthcare. An appendix of useful tools for evaluation of AI in healthcare companies as well as AI strategies in healthcare organizations is included.



MEDICAL INTELLIGENCE (TABLE OF CONTENTS)

I. INTRODUCTION

Artificial Intelligence: Basic Concepts

- Definitions
- Weak vs Strong or Narrow vs General AI
- SuperIntelligence and Singularity

Neuroscience and Artificial Intelligence

- Biomimicry of Brain by Machine
- Image Recognition and the Visual Cortex
- Machine Learning and the Brain
- Natural Language Processing and the Language Center
- Decision Support System and the Brain

II. THE EARLY YEARS OF ARTIFICIAL INTELLIGENCE & APPLICATION IN MEDICINE

A Brief History of Artificial Intelligence

- Alan Turing and the Turing Machine
- The Dartmouth Conference
- The AI Winters
- IBM Deep Blue and Watson

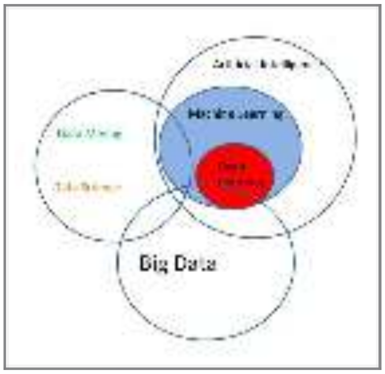
History of Artificial Intelligence in Medicine

- Rule-based Expert Systems and MYCIN
- Other Methods
- Reasons for Failure to Adoption

III. THE CURRENT ERA OF ARTIFICIAL INTELLIGENCE & APPLICATION IN MEDICINE

Introduction

The Current Era of Artificial Intelligence



- Algorithm
- Big Data
- Cognitive Computing
- Deep Learning

Other Key Concepts

- Natural Language Processing (NLP)
- Internet of Things (IoT)
- Computer Vision
- Robotics

Current Concepts of Artificial Intelligence in Medicine

Doctors and Machines

- How Doctors Think
- Comparing Doctors and Data Scientists

Healthcare Data and Databases

- The Conundrum of Health Care Data
- Healthcare Databases
- The Data-to-Intelligence Pyramid

Deep Learning and Healthcare

Current Applications of Artificial Intelligence in Medicine

- Decision Support and Hospital Monitoring
- Medical Imaging and Biomedical Diagnostics
- Precision Medicine and Drug Discovery
- Cloud Computing and Cybersecurity
- Digital Medicine and Wearable Technology
- Robotic Technology and Virtual Reality

Artificial Intelligence in Medicine in Selected Subspecialties

- Digital Medicine
- Cardiology (Adult and Pediatric)

- Oncology
- Pathology
- Radiology
- Ophthalmology
- Critical Care Medicine
- Psychiatry and Mental Health
- Pediatrics
- Genomic Medicine
- Infectious Disease
- Emergency Medicine
- Dermatology
- Endocrinology
- Population Health
- Global Health

Problems and Potential Solutions (Real Applications)

IV. THE FUTURE OF MEDICAL INTELLIGENCE

Introduction

The Future of Artificial Intelligence

- Augmented and Virtual Reality (Mixed Reality)
- Blockchain and Cybersecurity
- Cloud AI and Cognitive Architecture (Recursive Cortical Network)
- Deep Learning and New Types of Deep Learning (Deep Reinforcement Learning and Capsule Network)
- Internet of Everything (IoE) and Embedded AI (eAI)

Other Key Concepts

- Robotic Process Automation (RPA)
- Swarm Intelligence
- Quantum Computing
- Neuromorphic Computing
- Hypergraph Databases
- Brain-Computer Interface

The Future of Medical Intelligence

COMMENTARIES (500 WORDS/1 GRAPHIC)

- Precision Medicine and Pharmacogenomics (Altman)
- Artificial Intelligence in Medicine with Exponential Impact (Banifatemi)
- Precision Medicine and Big Data (Butte)
- The Informatics Consult Service (Callahan/Shah)
- Clinical Aspects of Representation Learning (Finn Catling)
- Public Database MIMIC III and Data Science (Celi)
- Educating the Dual Physician-Data Scientist for the Future (ACC)
- Multidisciplinary Datathons (ChenC)
- The Data Conundrum and Possible Solutions (ChenL)
- Internet of Things and Concept of Precision in Healthcare (Chou)
- Evolution of Clinical Informatics Towards Artificial Intelligence (Coeira)
- Human Soft Skills for Data Scientists (Cooper)
- Data Sharing for a Learning System (Cummings)
- Informatics in the Next Century for AI (Feaster)
- Experience with Cognitive Computing and IBM Watson (Gago)
- Cybersecurity and Blockchain (Hall)
- The Relationship Between CIO and the AI Team (Henderson)
- Capsule Network and Implication for Biomedicine (Hinton)
- The Concept of Chaos in Biomedicine (Holbrook)
- Deep and Reinforcement Learning and the Kaggle Phenomenon (Howard)
- Reinforcement Learning in the Clinical Setting (Komorowski)
- Future of Radiology and Other Medical Subspecialties (Langlotz)
- Clinician Adoption of AI (Maher)
- A New Artificial Intelligence Paradigm for Biomedicine (Marcus)
- Big Data in Healthcare (Mattison)
- Startup Strategies in AI in Medicine (Meyer)
- Wearable Technology and Data (Mohta)
- Math and Clinician (Moorman)
- Graph and Hypergraph Databases (Mousses)
- Ontology and its Use in AI (Musen)
- Google DeepMind in Biomedicine (Po)
- AI in Medicine and the FDA (Seals)
- Design Thinking in Biomedical Artificial Intelligence (Sherman)
- MYCIN and Early Experience with Artificial Intelligence in Medicine (Shortliffe)
- The Artificial Intelligence in Medicine Curriculum (Szolovits)
- Education of AI in Medicine for Physicians or Genetic Algorithms (Tal-Botzer)

- Artificial Intelligence as a Neuroscience (Taraman)
- Women and Data Science (Valentine)
- Interpretable or Explainable Artificial Intelligence (Wang)
- The Virtual Pediatric Intensive Care Unit Concept (Wetzel)

Others

- Legal Aspects of AI in Medicine and Healthcare (TBD)
- The One Hundred Year AI Study (AI100)(TBD)
- Ethical Issues of AI in Medicine

ARTIFICIAL INTELLIGENCE IN MEDICINE COMPENDIUM

KEY REFERENCES

- Books on Artificial Intelligence
- Books on Data and Artificial Intelligence in Biomedicine
- Journals and Journal Issues
- Articles
- Quotes
- Websites
- Movies/Videoclips
- People
- Companies

- Evaluation of Healthcare Organizations and Companies
- Elements for Successful Implementation of AI in Healthcare Organizations

GLOSSARY

TUTORIALS

- R
- Python
- MatLab

"Healthcare is an information industry that continues to think that it is a biological industry."
Laurence McMahon at the AAHC Thought Leadership Institute meeting, August, 2016

I. INTRODUCTION

On March 10th, 2016, Google's *AlphaGo* software made the Go game's 37th move as it competed against the best human Go champion Lee Sedol: this move was so astonishing in its ingenuity that Sedol felt compelled to leave the room to recover. This moment in which the computer or machine intelligence may have created an entirely novel Go strategy heralded yet another dawning of a new era in artificial intelligence (AI).

Major universities with AI departments as well as technology giants such as Google, IBM, Apple, Facebook, and Microsoft are all fervidly exploring real life applications for AI. Even though the advent of data science and machine learning has advanced information such as financial interactions and sports performance and promoted innovations such as the autonomous car and even a music album named *I AM AI* that is the first composed and produced by an AI, healthcare and medicine remain very much behind these other sectors in leveraging this new AI paradigm. Despite this, however, the recent major escalation of venture capital into healthcare and AI domain promulgated currently over 100 companies in AI in healthcare with an expectant $5-10 billion to be spent per year on AI in healthcare near 2020. Two major unicorns emerged in 2016: iCarbonX (from China) focuses on personalized health and Flatiron Health (based in the U.S.) promotes an oncology cloud platform for providers and life science companies.

Since the first article published in this AI in biomedicine domain in 1958 ([4]), there remains a relative paucity of published reports in medical journals and a congruent lack of interest amongst clinicians in applications of AI in medicine (AI and related topics consisted of 11,225/4,419,743 or 0.25% of total publications in medicine-related references). In 2016, there were only about 1,274 reports on AI applications in medicine out of a total of close to 408,000 articles found on PubMed, or a mere 0.31% (under search terms "artificial intelligence" and "medicine" with a partial list of more relevant subspecialties, see below). If one were to use a more specific AI terminology like "machine learning", the number of publications dwindles significantly to about 813 (or 0.20% of total publications); this is still a dramatic improvement from 2006, when there were only a total of 20 reports with the key term "machine learning" out of a total of 146,697 papers in medicine (or 0.01%).

[4] Rosenblatt F. The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain. *Psychol Rev* 1958; 65(6): 386-408.

Although a few journals remain stalwarts not entirely supportive of AI even now (*New England Journal of Medicine* has only published a few editorials) ([5]), others such as *Lancet, Annual Review of Medicine, JAMA,* and *British Medical Journal* as well as other journals with relatively high impact factors have recently published reports using artificial intelligence and machine learning (see below)([6]).

---

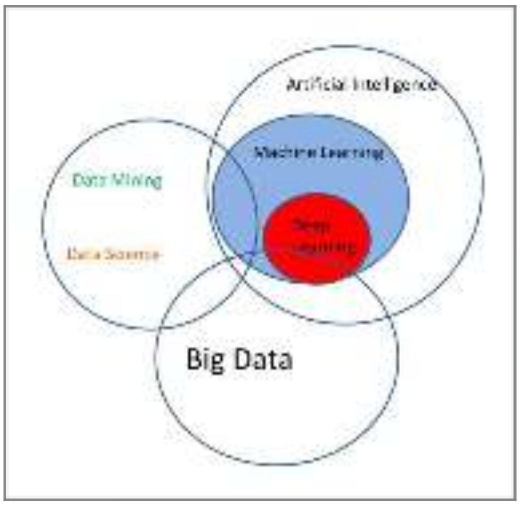
[5] Chen JH and Asch SM. Machine Learning and Prediction in Medicine- Beyond the Peak of Inflated Expectations. *N Engl J Med* 2017; 376(26): 2507-2509.

[6] Shapshay SM. Artificial Intelligence: The Future of Medicine? *JAMA Otolaryngology* 2014; 140(3): 191.

Artificial Intelligence: Basic Concepts

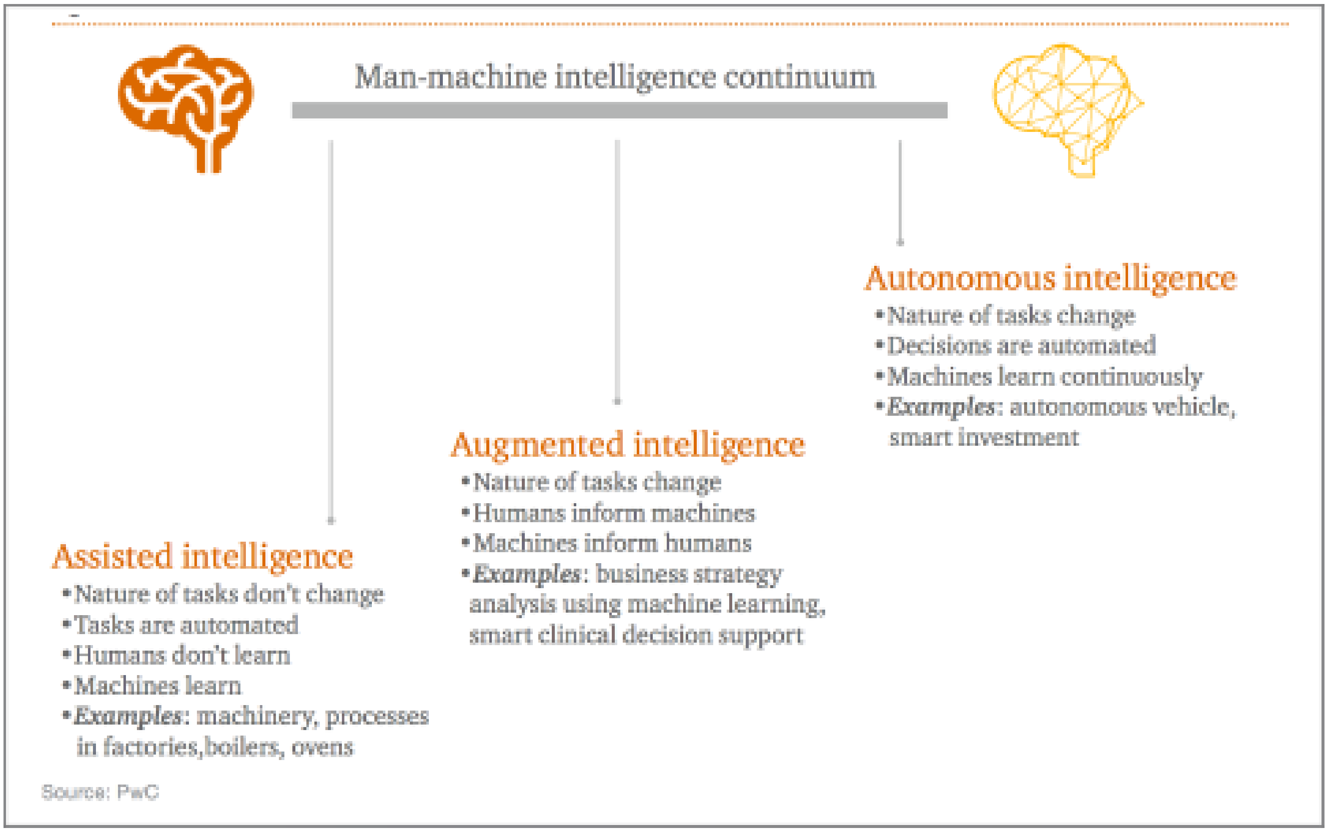
Intelligence can be defined as the ability to learn or understand or to deal with new situations or to apply knowledge or skills to manipulate one's environment. These definitions have interesting implications for artificial intelligence. Perhaps the best definition of artificial intelligence is the one conjured by the American cognitive scientist Marvin Minsky: the science of making machines do things that would require intelligence if done by man (woman).

Artificial intelligence can be categorized as weak vs strong: weak (or specific, narrow) AI pertains to AI technologies that are capable of performing specific tasks (like playing chess or *Jeopardy!*) and strong (or broad, general) AI, also called artificial general intelligence (or AGI), relates to machines that are capable of performing intellectual tasks that involve human elements of senses and reason. The public perception of artificial intelligence, however, continues to be that of the menacing robots that threaten mankind (such as HAL in *2001: A Space Odyssey* or the Terminator). Recently, this perception is modified to that of the more sophisticated and complex artificial intelligence-inspired but humanoid robots seen in the movies *Her* (2013) and *Ex Machina* (2015).



Machine learning (and its specific domain deep learning) are not synonymous with artificial intelligence but are rather types of AI methodology (see Figure). AI, however, does overlap with data science and data mining as well as big data.

Other AI methodologies can include cognitive computing and natural language processing (not in Figure). Cognitive computing (as exemplified by IBM's Watson cognitive computing platform) can involve a myriad of AI tools that simulates human thinking processes while natural language processing involves connecting human language with computer programmed understanding.



Artificial intelligence also consists of a man-machine intelligence continuum (see Figure) with three types of artificial intelligence: assisted, augmented, and autonomous.

The Swedish philosopher Nick Bostrom cautions the advent of a superintelligence that is essentially an intelligent agent that is superior to humans in intelligence ("an intellect that is much smarter than the best human brains in practically every field, including scientific creativity, general wisdom, and social skills")([7]). The futurist Ray Kurzweil similarly describes a technological singularity, a phenomenon in which the exponential increase in machine intelligence will supersede the human intelligence near the year 2045 ([8]).

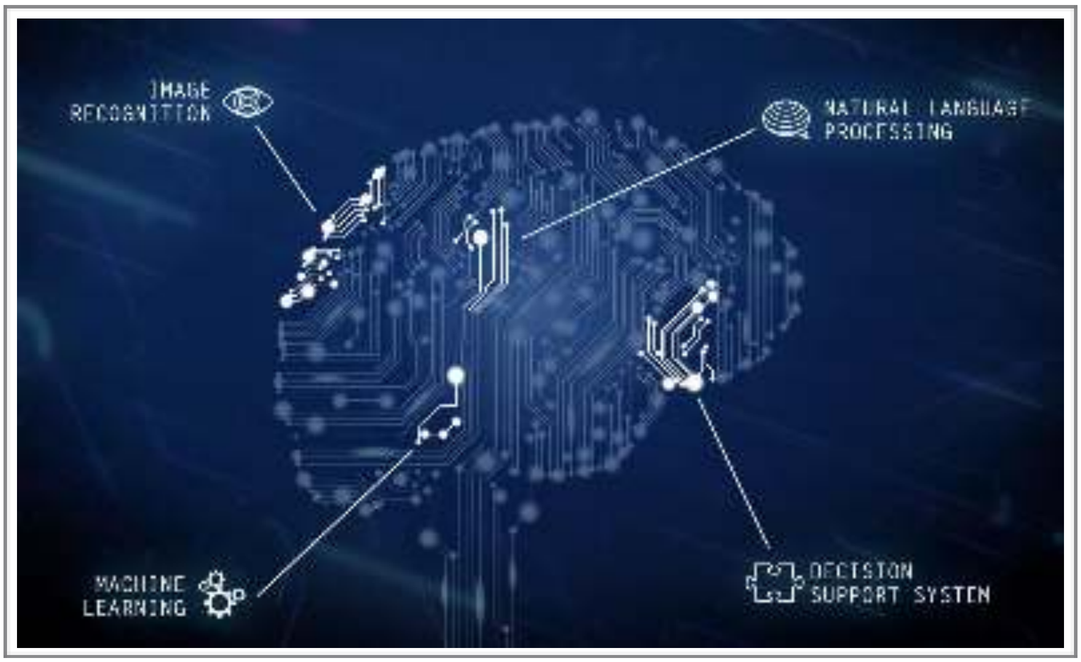
[7] Bostrom, N. *Superintelligence: Paths, Dangers, Strategies.* Oxford University Press, London, 2014.

[8] Kurzweil, R. *The Singularity is Near.* Viking Press, New York City, 2005.

Neuroscience and Artificial Intelligence

In Greek mythology, Icarus once yearned to fly like the bird by mimicking the bird but failed; instead, man eventually learned to fly by building planes and spaceships after attaining a deep understanding of the principles of aerodynamics. Similarly, one needs to build machine intelligence and artificial intelligence after a thorough comprehension of the brain and neuroscience ([9]). In short, innovative AI systems can be inspired by the brain.

The "doctor's brain" (see Figure) for day-to-day clinical work can be deconstructed by its myriad of functions and machine-equivalent capabilities. A cardiologist, for instance, will need to review medical images in the form of an electrocardiogram, echocardiogram, and perhaps an MRI or CT scan. He/she also needs to think about a patient's case after hearing about the history from the patient and/or family



members. Lastly, he/she needs to make complex diagnostic and therapeutic decisions.

Image recognition by the retina and visual cortex of the occipital lobe can be done by computer vision with interpretation coupled to machine learning. Thinking about the nuances of a particular clinical case is mimicked by machine learning. Language comprehension and speech by the brain's Broca's and Wernicke's areas of the left hemisphere can be performed by natural language processing. Lastly, the doctor's prefrontal lobe is part of the brain that helps to make challenging decisions and this process is mimicked by the computer's decision support system.

[9] Hassabis D, Kumaran D, Summerfield C et al. Neuroscience-Inspired Artificial Intelligence. *Neuron Review* 2017; 95(2):245-258.

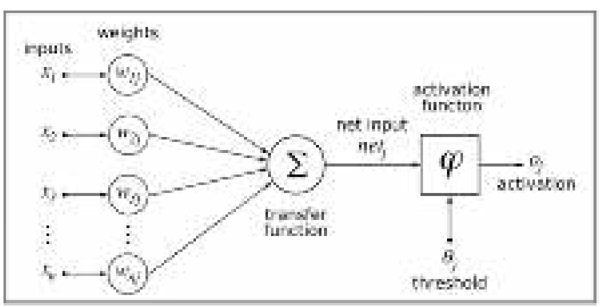
II. THE EARLY YEARS OF ARTIFICIAL INTELLIGENCE & APPLICATION IN MEDICINE

A Brief History of Artificial Intelligence

The history of artificial intelligence can be traced back to the study of logic delineated by the Greek philosopher Aristotle as he formulated a system of syllogisms for proper reasoning. The statistician Thomas Bayes and his framework for probability, the mathematician George Boole and his Boolean algebra, and the polymath Charles Babbage and his early digital programmable computer (the Analytic Machine) all contributed to the underpinnings of present day artificial intelligence.

It is the British mathematician and computer scientist Alan Turing, however, who would be considered the absolute progenitor of artificial intelligence with his pioneering works that included his theory of computation and his work on computing machines ([10])([11]). His most valuable contribution was his deciphering of the German Enigma machine during the second World War at Bletchley Park using machine intelligence (portrayed in the film *The Imitation Game*). The eponymous Turing Test is a test of machine AI's ability to pass as a human.

In 1956, mathematicians and scientists gathered at the seminal Dartmouth Conference and it is the proposal for this august gathering that the term "artificial intelligence" was coined by the Stanford computer scientist John McCarthy. This summer conference and its discussions is widely thought to be the birth of AI as an interdisciplinary field.

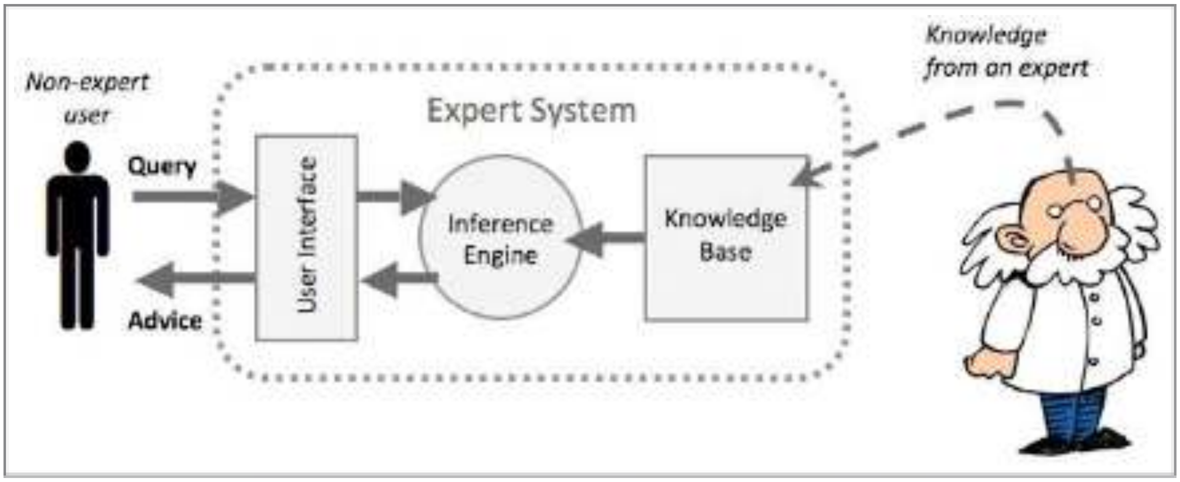


Around this time, a significant contribution was made by the American psychologist Frank Rosenblatt in 1958 in the form of the perceptron, a three layer structure (input, transfer function, and output) which was the early precursor of the artificial neural network and present day deep learning (see aforementioned reference).

[10] Turing AM. On Computable Numbers, with an Application to the Entscheidungsproblem. *Proceedings of the London Mathematical Society,* Series 2, 42 (1936-7), pp 230-265.

[11] Copeland, J. *The Essential Turing.* Oxford University Press, Oxford, 2004.

Following this early epoch of machine intelligence, two AI "winters" in the 1970's and then subsequently in the following decade occurred due to concomitant lofty expectations and suboptimal realities, resulting in an overall disappointing outlook on AI. Main shortcomings include the lack of a theory-to-use coupling as well as the inadequate integration of the existing AI techniques into workflows to achieve user support.



History of Artificial Intelligence in Medicine

Initial efforts in artificial intelligence and its application in medicine began in the 1960's and focused mainly on diagnosis and therapy. Among the best known early works on AI in medicine was the Stanford physician and biomedical informatician Edward Shortliffe's innovative heuristic programming project MYCIN. This pioneering work was a rule-based expert system (written in the Lisp programming language) that had if-then rules; these rules yielded certainty values that mimicked a human's expertise (such as recommended selection of antibiotics for various infectious diseases)([12]). The knowledge from a human expert was entered into a knowledge base, which in turn was connected to an inference engine (see Figure). The non-expert user then queries a user interface that was coupled to the inference engine. The advice was then given to the user via this user interface.

Although not successfully used in an actual clinical setting, MYCIN proved to be superior to human infectious disease experts when a comparison was performed. Of note, a precursor to MYCIN was the DENDRAL expert system which helped to identify unknown organic molecules.

[12] Shortliffe EH, David R, Axline SG et al. Computer-based Consultations in Clinical Therapeutics: Explanation and Rule Acquisition Capabilities of the MYCIN System. *Comput Biomed Res* 1975; 8(4): 303-320.

During this early era of AI in medicine, several academic centers in the United States that focused on this area included Stanford, MIT, Rutgers, and Pittsburgh as well as a few centers in Europe. The MIT professor of Computer Science and Engineering Peter Szolovits edited a book on AI in Medicine in 1982 that consisted of a collection of papers on various topics in this domain ([13]). Szolovits later organized a Medical Artificial Intelligence course at MIT in 2003 that was one of the first organized educational efforts on this burgeoning topic.

Other AI methodologies used in medicine other than the traditional expert systems included methodologies such as fuzzy logic and neural networks ([14]). Fuzzy logic deals with degrees of truth (vs the dichotomous Boolean logic of true or false) and is therefore particularly well suited for biological systems with objective parameters (such as heart rate and blood pressure)([15]). Neural network is a processing paradigm that is inspired by the brain; this methodology was applied to various clinical situations such as clinical diagnosis and medical images as well as the critical care setting ([16]).

In conclusion, while this early period was initially characterized by knowledge engineering for medical expert systems, machine learning and data mining were AI methodologies in the later part of this early era. Artificial intelligence and its failed adoption in medicine during this early period was due to lack of favorable work flow logistics and integration to accommodate the clinicians and also due to expectations that were unrealistically high but never met.

---

[13] Szolovits P. *Artificial Intelligence in Medicine.* Westview Press Inc, Boulder, Colorado, 1982.

[14] Hanson CW and Marshall BE. Artificial Intelligence Applications in the Intensive Care Unit. *Crit Care Med* 2001; 29:427-435.

[15] Ramesh AN, Kambhampati C, Monson JR et al. Artificial Intelligence in Medicine. *Annals of the Royal College of Surgeons of England* 2004; 86(5): 334-338.

[16] Yardimci A. A Survey on the Use of Soft Computing Methods in Medicine. In *Proceedings of the 17th International Conference on Artificial Neural Networks,* Porto, Portugal, 69-79.

III. THE CURRENT ERA OF ARTIFICIAL INTELLIGENCE & APPLICATION IN MEDICINE

Introduction

The data mining and machine learning focus in the 1990's then slowly revived the field of AI and this era was best symbolized by IBM's supercomputer Deep Blue, which defeated the reigning world chess champion Gary Kasparov in 1997. Another IBM supercomputer, Watson (named after its first CEO Thomas Watson), with access to over 200 million pages of content and developed in IBM's DeepQA (question and answer) project, easily defeated the human champions Ken Jennings and Brad Rutter on February 14, 2011, on the game show *Jeopardy!.* In a similarly dominant fashion, the AlphaGo program of DeepMind easily defeated the human Go champion Lee Sedol in March 2016, thus heralding a new era of AI with deep learning.

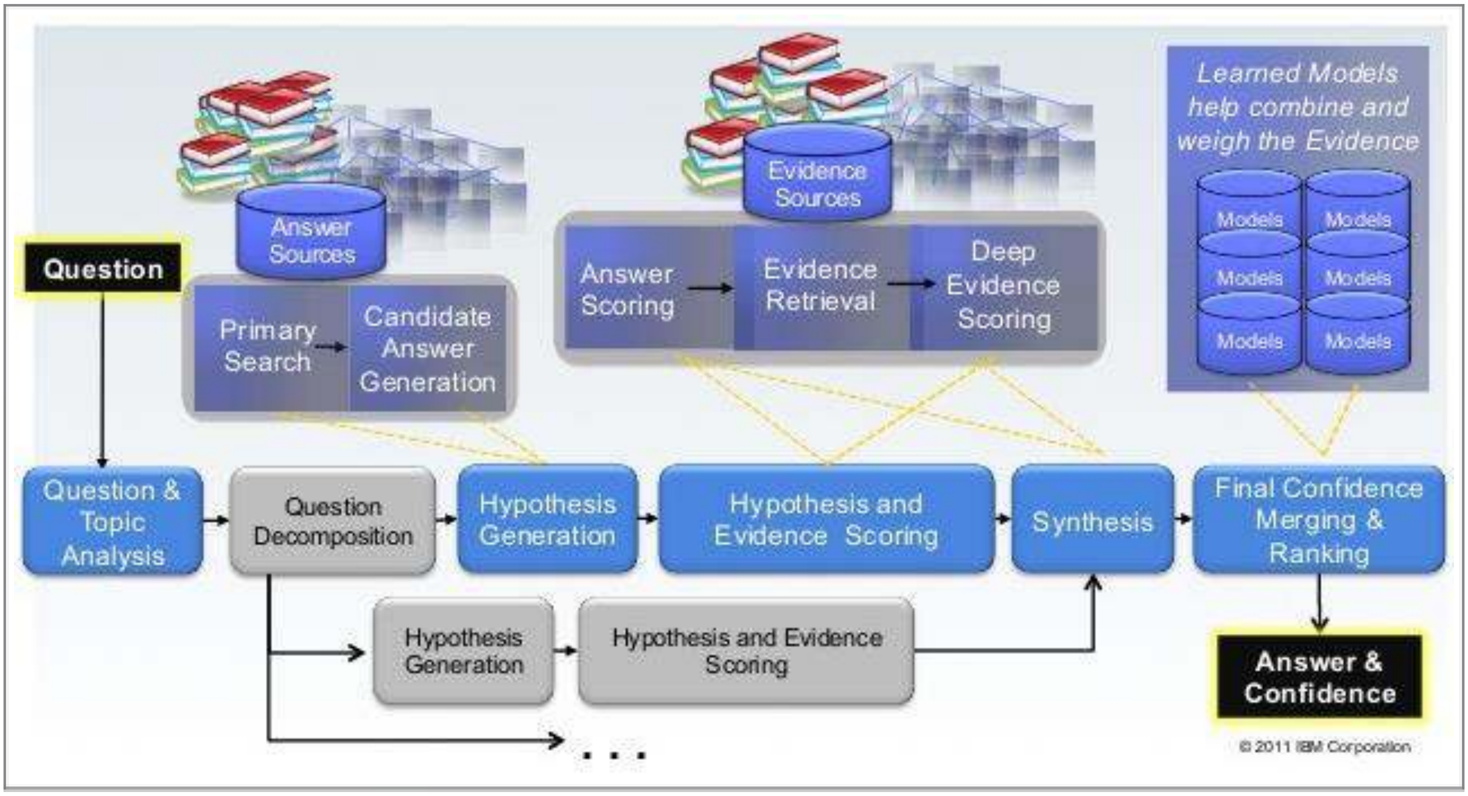
The recent advent of an AI "trinity" that consisted of: 1) the increasingly large volumes of available data that requires new computational methodologies (or simply "Big Data"), 2) the escalating capability of computational power (with faster, cheaper, and more powerful parallel processing that defied Moore's Law) and cloud computing (with nearly infinite storage), and 3) the emergence of machine and deep learning with its variants have together promulgated this new dawn of AI.

The Current Era of Artificial Intelligence (The "A, B, C, and D" of Current AI)

Algorithms. The advent of complex and efficient algorithms (sets of steps to accomplish certain tasks) that are available for not only calculations and data processing but also automated reasoning has advanced the capabilities of machine intelligence. Examples of complex algorithms that are in current use include Pixar's coloring of 3D characters in virtual space (rendering algorithm) and NASA's operations of the solar panels on the international space station (optimization algorithm).

Big Data. Data have escalated in a myriad of ways to the point that traditional data processing applications are no longer adequate. The four "V"s of big data often discussed are: 1) volume (over 40 zettabytes, or the equivalent of 40 trillion gigabytes, are expected to be in existence by 2020 with internet of things accelerating this growth), 2) variety (videos, wearable technology, tweets, and structured vs unstructured types of data can create a digital chaos), 3) velocity (speed data is accessed such as with streaming data and over 20 billion network connections by the end of this year), and 4) veracity (uncertainty of data is not only costly but leads to inaccurate conclusions). Additional "V"s in big data include: value, visualization, and variability.

Cognitive Computing. Cognitive computing uses machine learning, pattern recognition, and natural language processing (NLP) as well as other AI tools to mimic the human brain and its self-learning capability. The IBM supercomputer Watson with its victory in the game show *Jeopardy!* against human champions in 2011 heralded the era of cognitive computing with its potent NLP and knowledge representation and reasoning capabilities along with machine learning (see Figure)([17]). The



supercomputer can scan 40 million documents in 15 seconds.

There is sometimes understandable confusion between AI and cognitive computing. While AI does not intentionally mimic human thought processes, cognitive computing with its origin in cognitive science, does attempt to simulate the human problem-solving process in a computerized model via AI tools such as machine learning, neural networks, and natural language processing as well as sentiment analysis and contextual awareness. While the present day virtual assistants are pre-programmed collection of responses, a cognitive system can yield a more thoughtful "human" response in the near future.

[17] Chen Y, Argentinis E, and Weber G. IBM Watson: How Cognitive Computing Can Be Applied to Big Data Challenges in Life and Science Research. *Clinical Therapeutics* 2016; 38(4): 688-701)

Machine Learning. Machine learning is an increasingly popular sub-discipline of AI and focuses on big data. In machine learning, a computer uses algorithms to find patterns in data. The sophisticated algorithms are used to interpret data (from a "training set") with the use of classifiers (features or attributes that are used to classify the subjects in a process called feature extraction) in order to make predictions (from an initial "test set" first followed by new datasets). In other words, the features are predictor variables with labeled outcomes. In short, the four steps of machine learning are: data pre-processing, feature extraction, machine learning algorithm, and predictive model as the last step (see

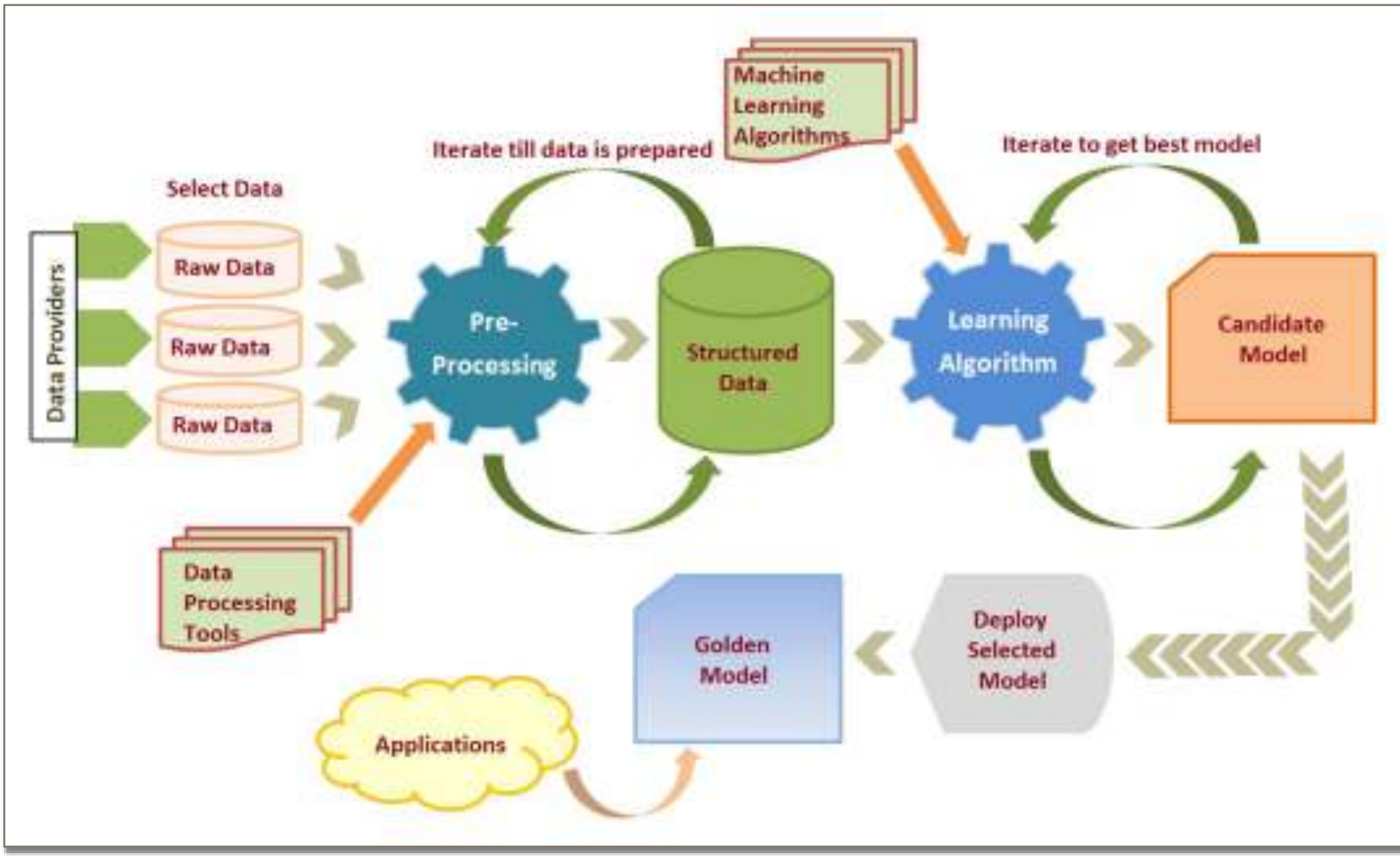
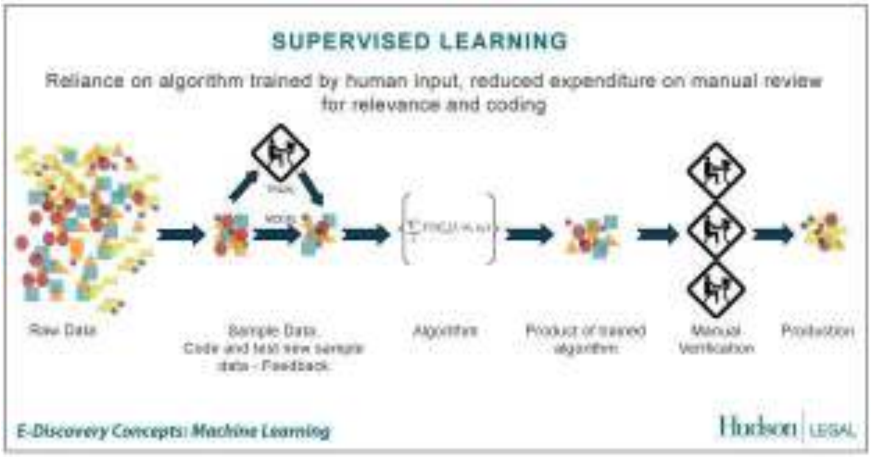


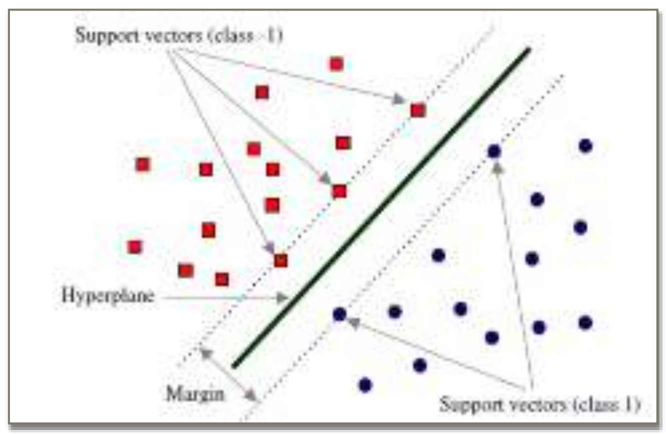
Figure).



Machine learning is usually categorized into three types of learning:

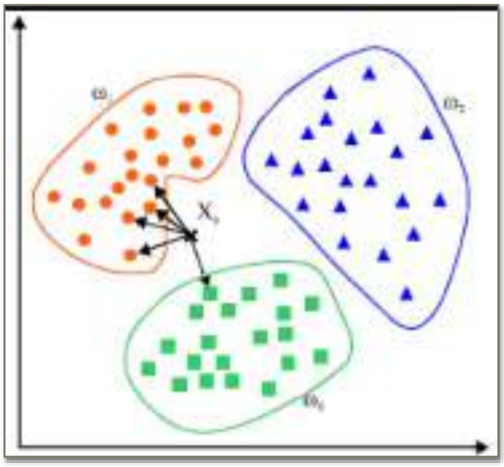
First, supervised learning take raw data and use an algorithm to predict the outcome based on a prior training set of data that are labeled.
These supervised learning methodologies lead to classification and regression. Classification leads to categorization of output variables whereas regression leads to numerical representation of output variables.

These supervised learning methodologies include: support vector machines (SVM), naive Bayesian classifiers, *k*-nearest neighbor (*k*-NN), linear and logistic regression, and decision trees methods (like random forest).



Support vector machines (SVMs). This machine learning methodology is good for complex nonlinear relationships by creating an optimal hyperplane in a high-dimensional space that represents the largest separation between the two classes (see Figure).
This methodology is often considered one of the most robust and accurate algorithms for classification especially when the number of features are high (compared to the number of data points) and is frequently used for pattern recognition especially when the training set is small (larger training sets may be better suited for deep learning).

Naive Bayes classifiers. Based on the Bayesian theorem (with its concept of prior probability which selects the outcome with the highest probability), this supervised learning methodology is relatively fast and also good when the dimensionality of the inputs is high. This methodology is well suited for real time prediction, text classification/spam filtering, and recommendation system.



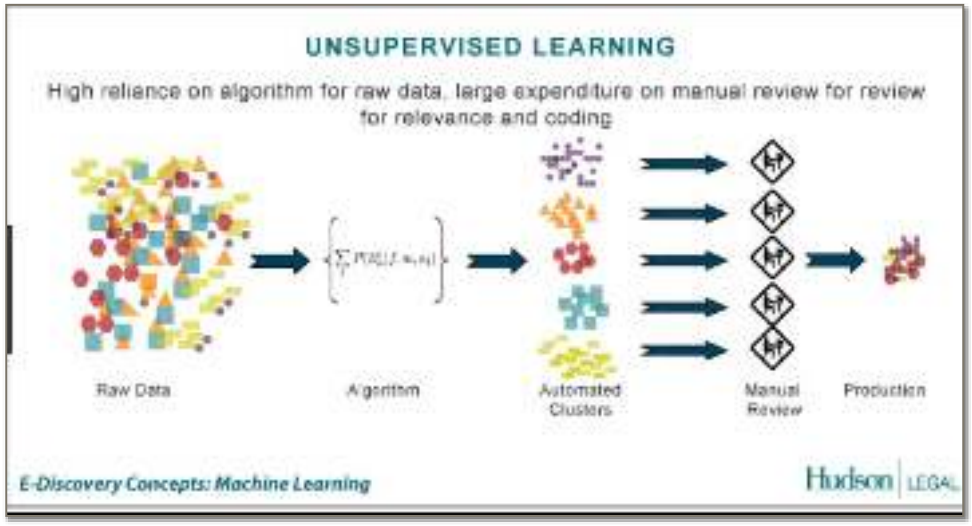
k-nearest neighbor (*k*-NN). This supervised learning algorithm is used for both classification as well as regression and identifies the number of nearest neighbors of any element. This methodology is considered a “lazy” learning that is instance-based. This methodology is well suited for data sets in which there is no prior knowledge about the distribution of the data such as text mining or categorization as well as stock market trends and forecasting.

Linear regression. This machine learning method (derived from statistics) delineates the relationship between two continuous variables (“x” is the independent variable whereas “y” is the dependent variable). This regression is termed “simple” when there is a single input variable and “multiple" when there are multiple input variables.

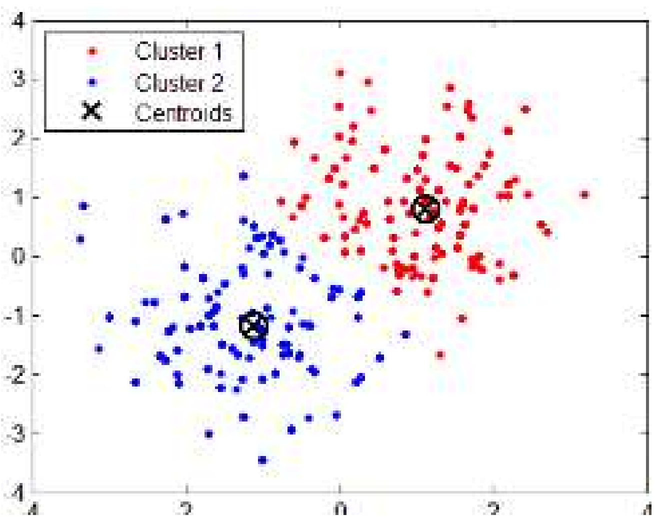
Logistic regression. This is the adaptation of the aforementioned linear regression to classification. It suffers the same problems as linear regression as both techniques are too simplistic for complex relationships between variables and have a tendency to “overfit”- a phenomenon in which a model is modeling the training data too well as to affect the performance of the model on new data.

Random forest. This decision-tree supervised learning methodology is often used due to its simplicity and flexibility as it can perform both classification as well as regression analyses. The forest is essentially an ensemble of many decision trees that collectively and yield an accurate prediction and therefore avoids overfitting. One major limitation of random forest is its slow nature (with large number of trees) for real-time predictions. This methodology is well suited for detecting fraud and stock projections.

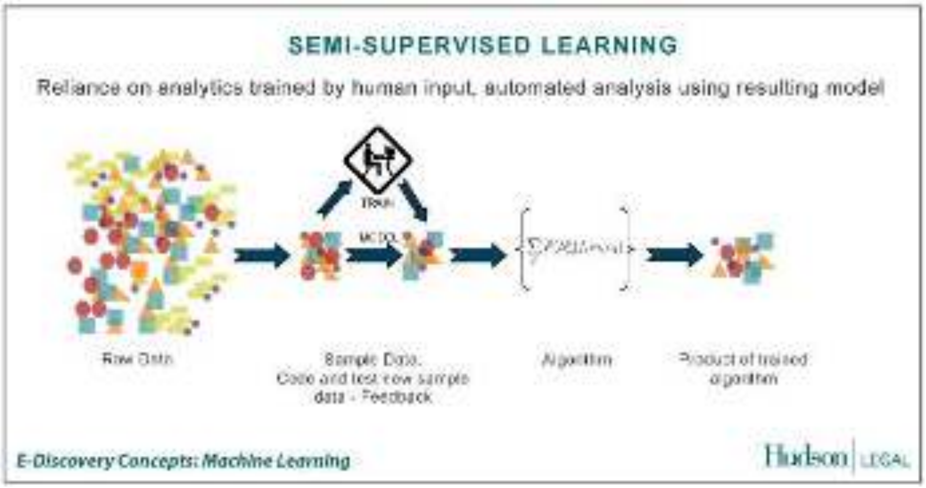
Second, unsupervised learning take unlabeled data and use algorithms to predict patterns or groupings in the raw data set. These unsupervised learning methodologies lead to clustering or association. Other questions unsupervised learning can answer include segmentation and dimension reduction.



These clustering methodologies include: *k*-means clustering and association methodologies can include apriori algorithm.



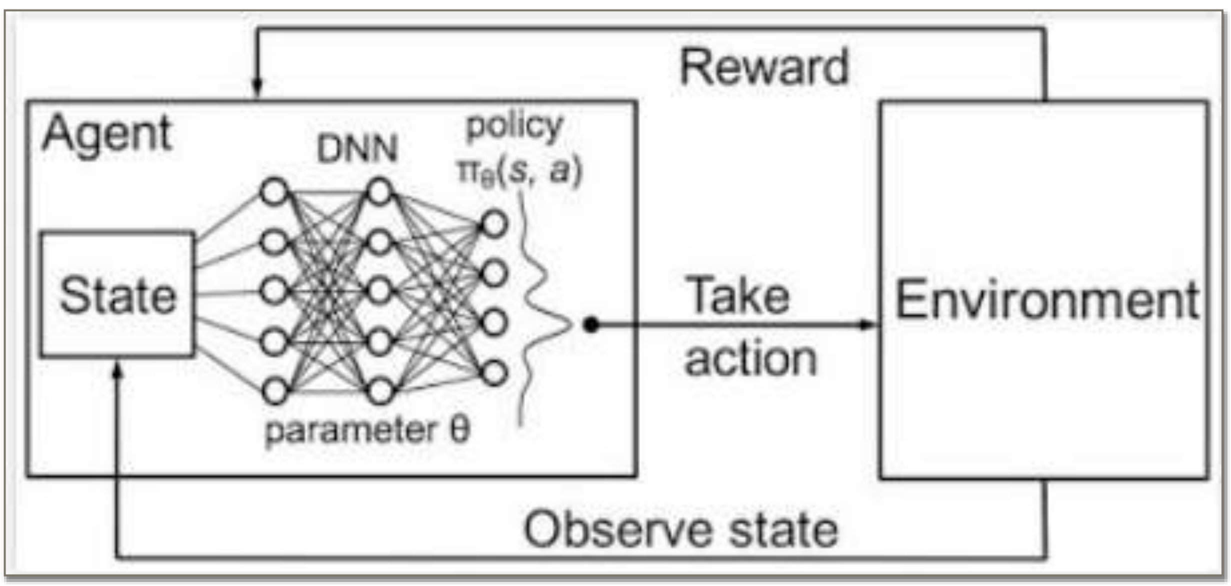
K-means clustering. This is a commonly used unsupervised learning that has an algorithm used to find clusters or groups in the data with variable K number of groups that form organically based on similarity (see Figure). The end result of the process yields cluster centroids of the clusters and the labels for the training data.



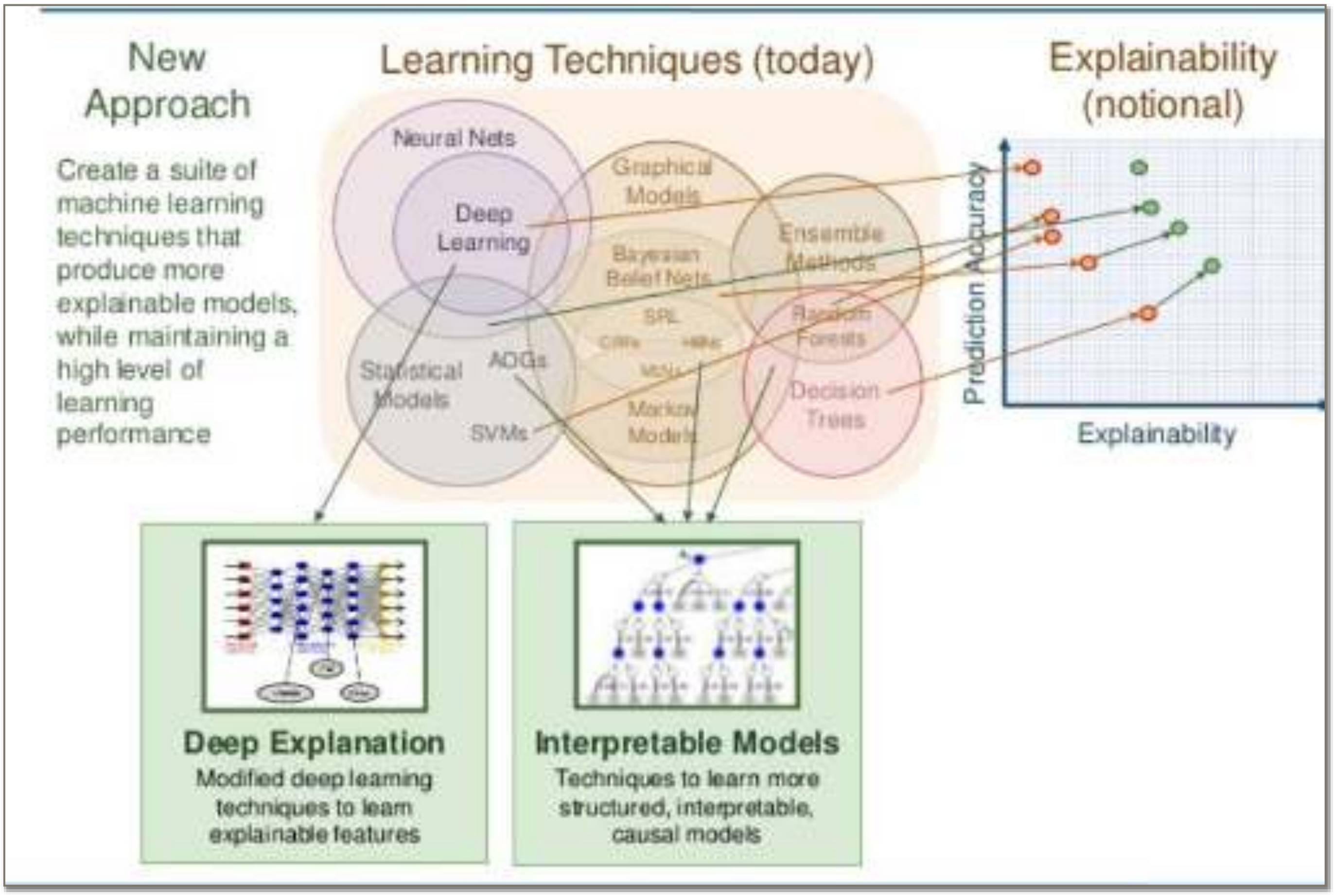
Recent hybrid techniques such as semi-supervised learning can be used with less labeled data than that required for supervised learning. These methodologies can therefore be trained on a mixture of labeled and unlabeled data. The introduction of unlabeled data may reduce human bias and improve accuracy of the final model.

In addition to the aforementioned supervised and unsupervised learning, a third type of learning is reinforcement learning. In this type of learning, the model finds the optimal method to achieve the most desirable outcome analogous to humans attempting to attain the highest score in a game (see Figure). In other words, there is a positive and negative feedback to the solution of the algorithm so reinforcement learning is well suited for decision process.

Reinforcement learning is the methodology that AlphaGo utilized in its defeat of the human Go champion and may be an asset for biomedicine as it is designed to make decisions in an uncertain environment.



There are several limitations with machine learning. A common issue with machine learning resides in its "black box" characteristic- for those who are not data scientists, it is difficult to understand the data science in the machine learning process (see Figure)([18]). Some of the higher prediction accuracy machine learning methodologies (deep learning, random forest, support vector machines, etc) have the least explainability whereas others (Bayesian belief nets, decision tress) have more explainability (but lower prediction accuracy). There is an ongoing effort to elevate explainability in the form of "explainable AI or

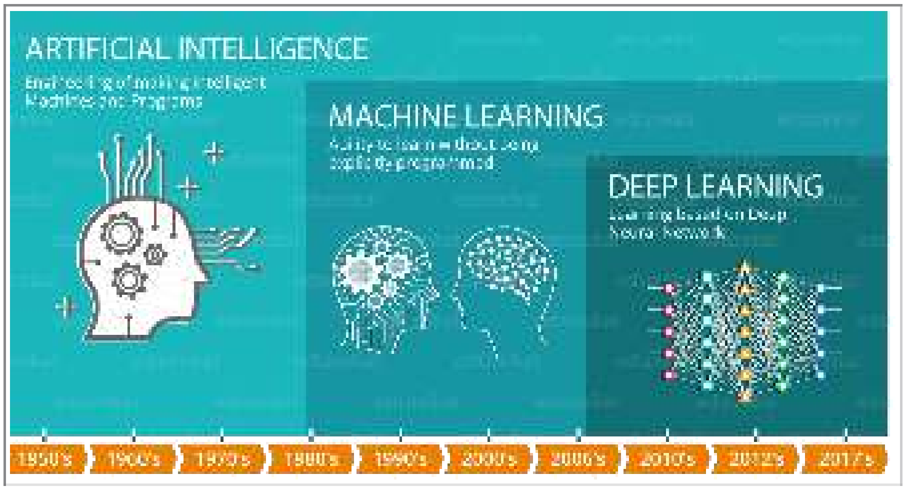


XAI") while maintaining (or even increasing) prediction accuracy with a new suite of techniques.

Another potential problem is overfitting: this phenomenon occurs with small studies that have concomitantly small number of attributes so the choice of attributes is critical for ideal machine learning. Circularity occurs when the classifier used in machine learning is validated by using the same data that were used to develop the classifier in the first place. Finally, validation of the predictive model is vital to machine learning and is often not optimal.

[18] Gunning, David. Talk at DARPA. August 11, 2016.

Deep Learning. A type of machine learning that was inspired by the brain with its neurons and intricate synaptic interconnections is termed neural network. These networks have layers of connections and data propagation directions (forward and backward). Back propagation is a mechanism in which the neural network can "learn": the difference between the output and the desired output is used to calculate a modification to achieve the desired outcome. The individual neurons can be assigned a weighting to the input signal and an accompanying probability vector.



In 2012, the team from University of Toronto used a deep learning algorithm with 650,000 neurons and 5 convolutional layers to reduce the error rate in half during a computer vision challenge ([19]). Andrew Ng of Stanford and Google and others synthesized huge neural networks by increasing the number of layers and neurons to enable large data sets to be trained to promulgate deep learning ([20])([21])([22]).

With open source software tools such as TensorFlow from Google, powerful supercomputers such as the NVIDIA DGX-1™, and quantum computing processors, deep learning is an exciting but also mysterious new extension of machine learning. Current applications of deep learning include speech recognition and natural language processing, computer vision with visual object recognition and detection, and autonomous vehicle driving.

---

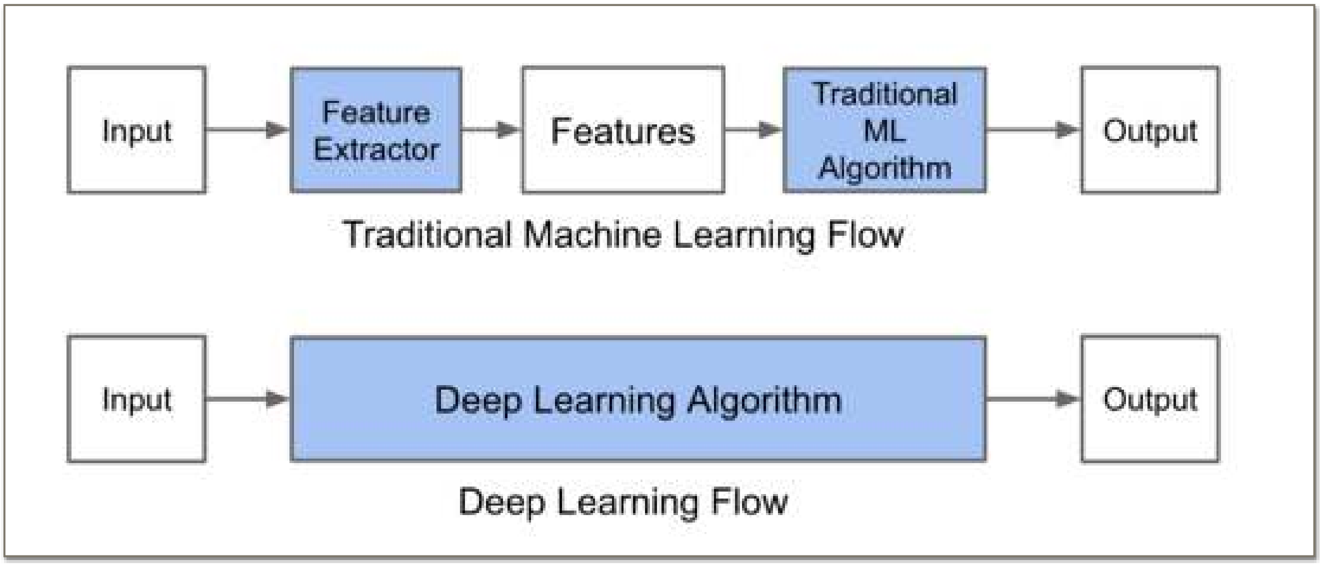
[19] Krizhevsky A, Sututskever I, and Hinto GE. *ImageNet Classification withDeep Convolutional Neural Networks.* Vol 1. La Jolla, CA: Neural Information Processing Systems Foundation Inc; 2012:4.

[20] LeCun Y, Bengio Y, and Hinton G. Deep Learning. *Nature* 2015; 521: 436-444.

[21] Porter J (ed). *Deep Learning: Fundamentals, Methods, and Applications.* Nova Science Publishers, New York, 2016.

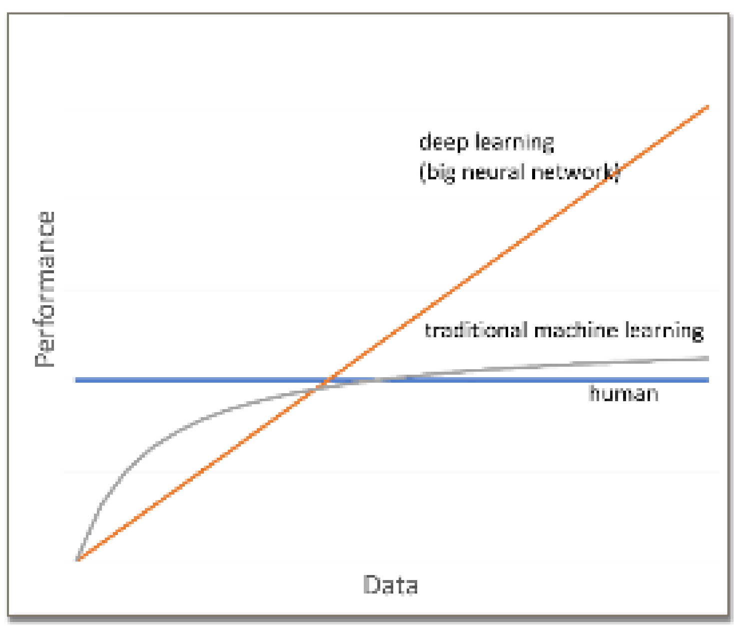
[22] Arel I, Rose DC, and Kanowski TP. Deep Machine Learning- A New Frontier in Artificial Intelligence Research. *IEEE Computational Intelligence Magazine* 2010; 1556-603X (13-18).

Whereas traditional machine learning flow has feature extraction followed by machine learning algorithm that leads to output, deep learning flow involves an artificial neural network that can combine



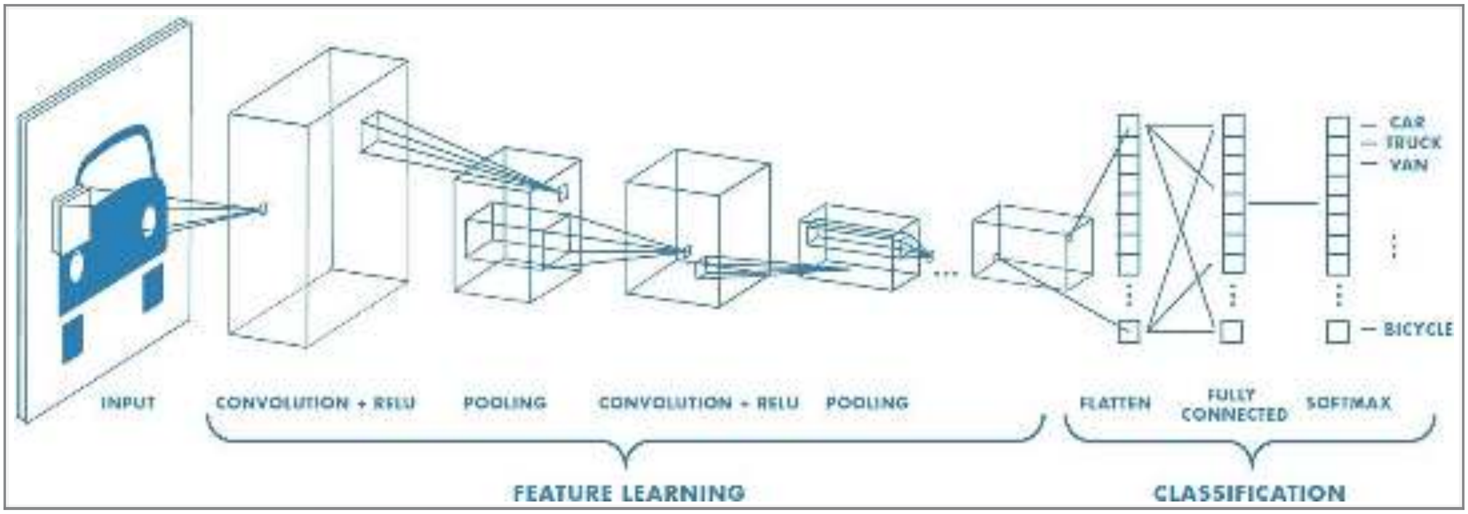
feature extraction with the classification as one step (see Figure).

Machine learning, compared to deep learning, is relatively easy to train and test but its performance is dependent upon its features and is limited even with increasing volume of data (see Figure). On the other hand, while deep learning can learn high-level features representation, it does require large amounts of data for training ("big data") and can be expensive from a computation usage perspective. In addition, deep learning are more difficult to comprehend as the algorithms are largely self-directed.



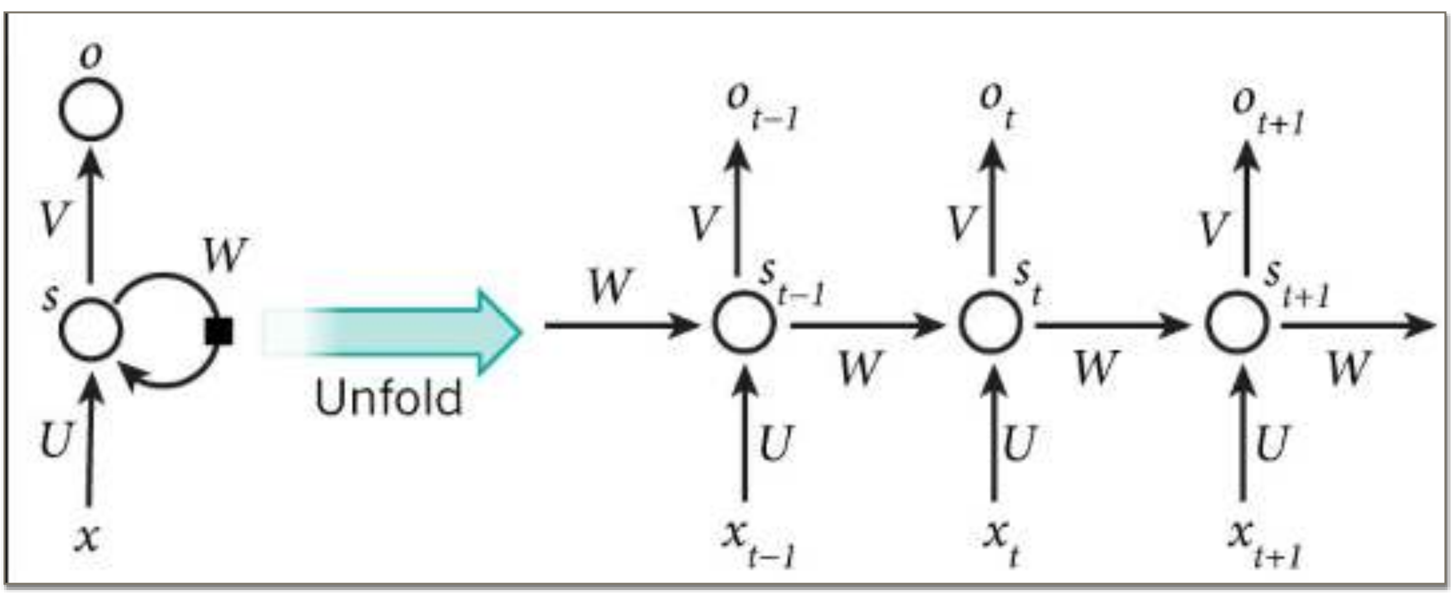
There are two branches in deep neural networks (DNN):

First, convolutional neural network (CNN) is a class of neural networks that consists of one or more three-dimensional "convolutional layers" inspired by how the human primary visual cortex functions. CNN is particularly good at image recognition and classification but has difficulty if images have



alterations like rotation or other orientation (see Figure).

A recurrent neural network (RNN) is the other branch of deep neural network that consists of one or more neural networks in which there is a feedback loop (next state depends on the prior state). RNN is therefore able to recall a memory as there are longer term dependencies compared to CNN. This type of deep learning is used therefore for repeating the same task for sequential information that are dependent of each other (like time series in the ICU setting or stock market analyses as well as language generation



or translation). There is also a hybrid "CNN-RNN" model that has potential in biomedical data.

Other Key Concepts

Natural Language Processing (NLP). This AI methodology allows the computer to understand spoken as well as written human language through specific set of techniques such as parsing, which is the analysis of words in the sentence for grammar followed by a process of arranging the words that shows the proper relationship between these words. The two components of NLP are natural language understanding (NLU) and natural language generation (NLG); NLU is usually considered the more difficult component. In short, NLP is the intersection of AI and linguistics.

Internet of Things (IoT). The recent explosion of physical devices and embedded sensors from equipment, buildings, vehicles, and appliances as well as wearable devices has lead to network connectivity which will enable all these devices to collect and exchange data. An IoT platform enables connectivity for these devices that are capable of being part of the network. The method of communication is radio-frequency identification (RFID) but can also be wireless, quick response (QR) codes, or other sensor technologies. The connected devices together can form ambient type of intelligence. IoT also needs to be coupled to the cloud.

Computer Vision. Machines gaining insight from digital images or from videos by use of artificial intelligence techniques such as deep learning for image interpretation and analysis. Tasks that can be accomplished with computer vision include: motion analysis, image restoration, and object recognition; the latter best done with convolutional neural networks. There has been much publicity recently on facial recognition using computer vision.

Robotics. This discipline is utilization of artificial intelligence and engineering to involve the conceptualization and design as well as operation of robots. The interdisciplinary science of robotics includes electrical and mechanical engineering along with computer science. The recent trend in robotics is to yield robots that are more humanlike and less mechanical. Finally, the field of robotics has

Asimov's Three Laws of Robotics

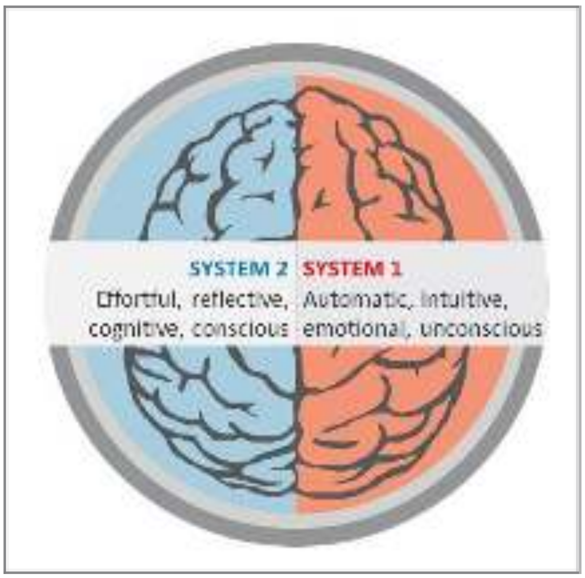
1. A robot may not injure a human being, or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence, as long as such protection does not conflict with the First or Second Law.

converged with others to now include avatars and virtual assistants.

Current Concepts of Artificial Intelligence in Medicine

Doctors and Machines

How Doctors Think. In Jerome Groopman's How Doctors Think ([23]), he aptly described several deficiencies in the way physicians think. One such mechanism is confirmation bias, which is the tendency for physicians to search for information that confirms one's preexisting hypothesis. In Sherlock Holme's parlance: "It is a capital mistake to theorize before one has data. Insensibly one begins to twist facts to suit theories, instead of theories to suit facts." Another example of cognitive error is the availability heuristic or an intellectual shortcut that relies on immediate recall when evaluating a situation. The myriad of human biases and heuristics can potentially be neutralized with an AI-supported strategy in decision-making process.



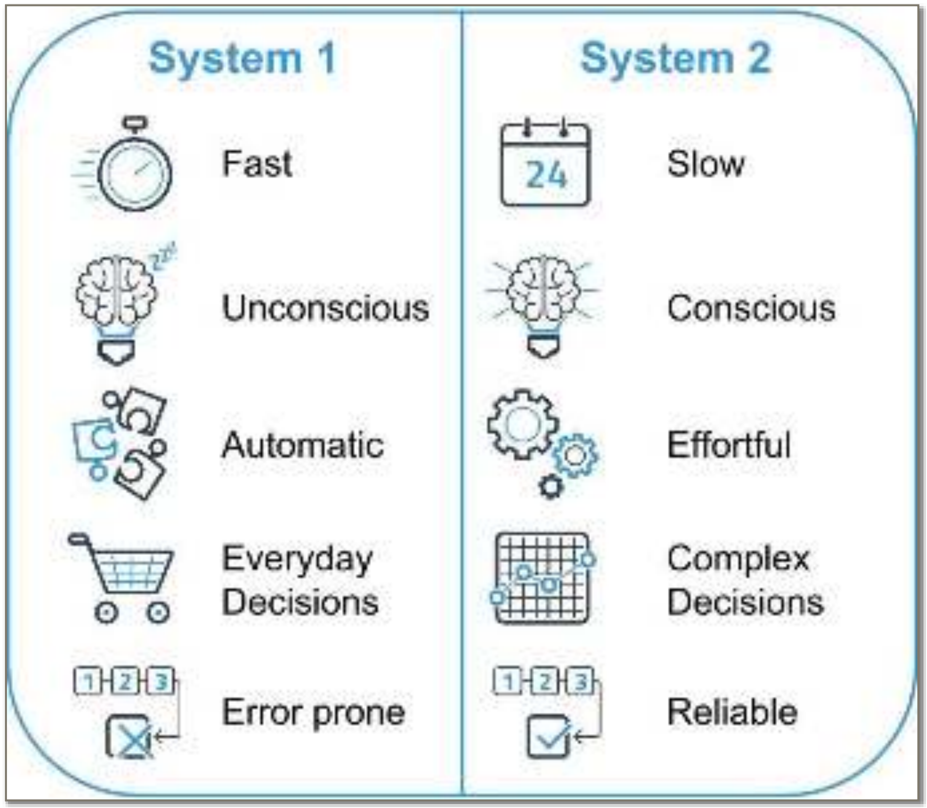
Comparing Doctors and Data Scientists. Daniel Kahneman, the Nobel Prize-winning psychologist noted for his work on decision making, described System 1 vs System 2 thinking (fast and experiential vs slow and analytical, respectively)(see Figure)([24]). This dichotomy conveniently delineates some of the key differences between clinicians (prone to System 1 thinking) and data scientists (with their affinity for System 2 thinking).

For example, physicians often rely on a fast intuition-based "System I" thinking that is based on experience and accumulated judgment. Data scientists, on the other hand, more frequently approach problems with slower and more logical progressive thinking that is rationality-based "System 2" thinking.

[23] Groopman J. *How Doctors Think.* Houghton Mifflin, Boston, 2007.

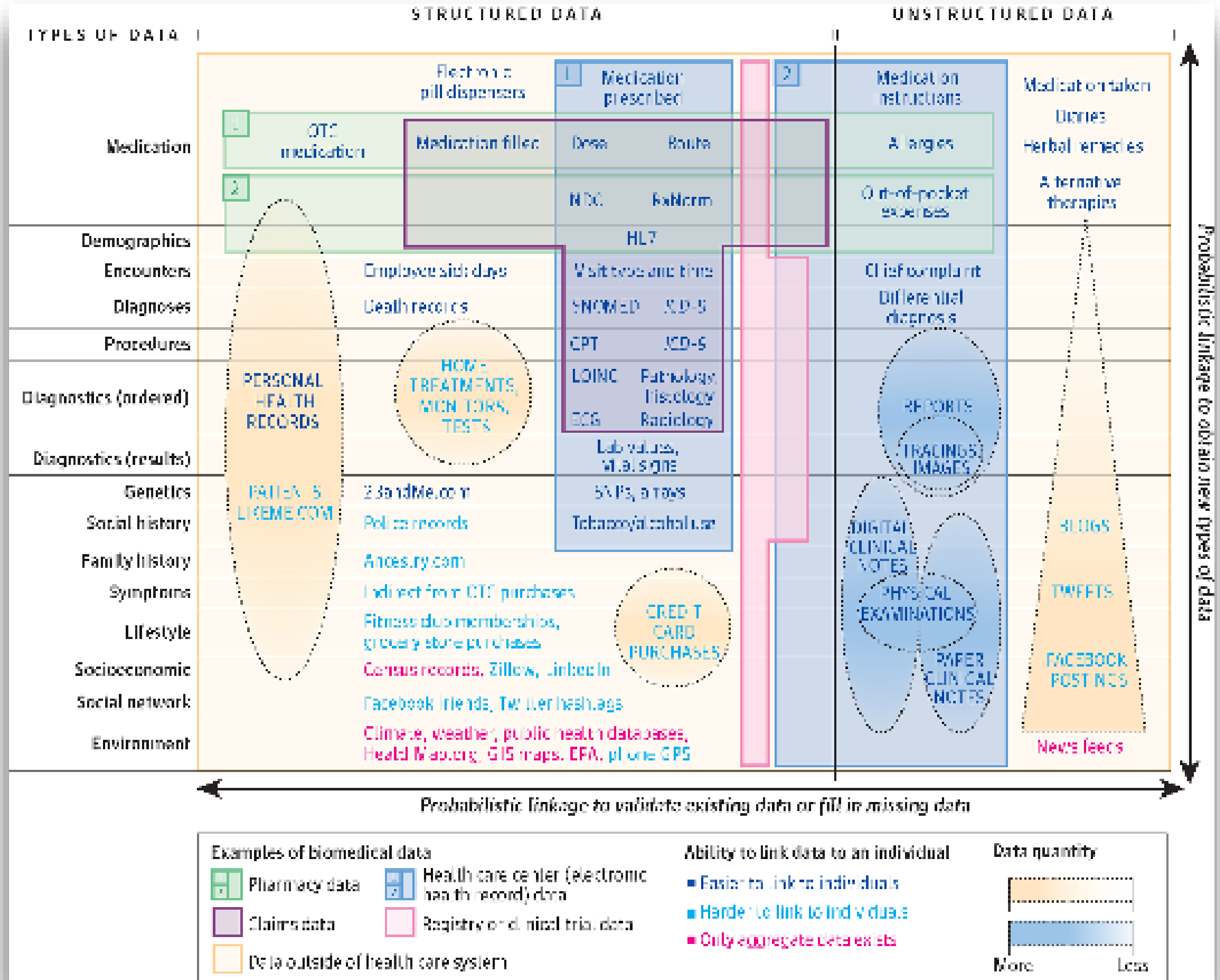
[24] Kahneman D. *Thinking, Fast and Slow.* Farrar, Straus, and Giroux, New York, 2011.

Similarly, the partnership between the inspector Sherlock Holmes and Dr. Watson describes their two predominantly different systems or "brains" for investigative work: the former (system 2) is more logical and objective (albeit cognitively more costly) while the latter (system 1) is more emotional and subjective (faster but with inherent biases and fallacies). Medicine ideally should perhaps incorporate both types of thinking and individualize decisions based on how much of either type is appropriate. This strategy will minimize the pitfalls in diagnosis and treatment due to inherent heuristics and biases in clinicians ([25]).



[25] Klein JG. Five Pitfalls in Decisions about Diagnosis and Prescribing. *BMJ* 2005; 330: 781-783.

Healthcare Data and Databases



The Conundrum of Healthcare Data. The current imbroglio in health care data is highlighted by an escalating volume of unstructured, heterogeneous medical data with little embedded predictive analytics or machine learning (see Figure)([26])([27]).

---

[26] Chang AC et al. Artificial Intelligence in Pediatric Cardiology: An Innovative Transformation in Patient Care, Clinical Research, and Medical Education. *Cong Card Today* 2012; 10: 1-12.
[27] Roski J et al. Creating Value in Health Care Through Big Data: Opportunities and Policy Implications. *Health Affairs* 2014; 33(7): 1115-1122.

The complex portfolio of health care data includes not only electronic medical records (patient encounters, vital signs, laboratory results, prescriptions, etc.) but also advanced imaging studies (such as MRI, CT scans, and echocardiograms and angiograms)([28]). In addition, it is estimated that about 80% of health care data is unstructured ([29]). Lastly, current estimate of health care data volume is above 150 exabytes in volume and escalating rapidly ([30]).
Despite the large volume, variety, and velocity of big data in biomedicine, there is little dividend in the form of information from this health care big data ([31])([32]). Yet, there are opportunities for utilizing health care big data to reduce costs: high-cost patients, readmissions, triage, decompensation, adverse events, and treatment optimization ([33]).

This situation will soon be far more complex and daunting with the advent of data "tsunamis": genomic data (as a result of the high throughput next generation sequencing)([34]) and physiologic data (from home monitoring and wearable physiologic devices)([35]).

---

[28] Weil AR. Big Data in Health: A New Era for Research and Patient Care. *Health Affairs* 2014; 33:1110.

[29] Healthcare Content Management White Paper: Unstructured Data in Electronic Health Record (HER) Systems: Challenges and Solutions. October, 2013. www.datamark.net.

[30] Hughes G. How Big is "Big Data" in Healthcare? *SAS Blogs*. October 11, 2011.

[31] Jee K et al. Potentiality of Big Data in the Medical Sector: Focus on How to Reshape the Healthcare System. *Healthc Infrom Res* 2013; 19(2): 79-85.

[32] Schneeweiss S. Learning from Big Health Care Data. *N Eng J Med* 2014; 370: 2161-2163.

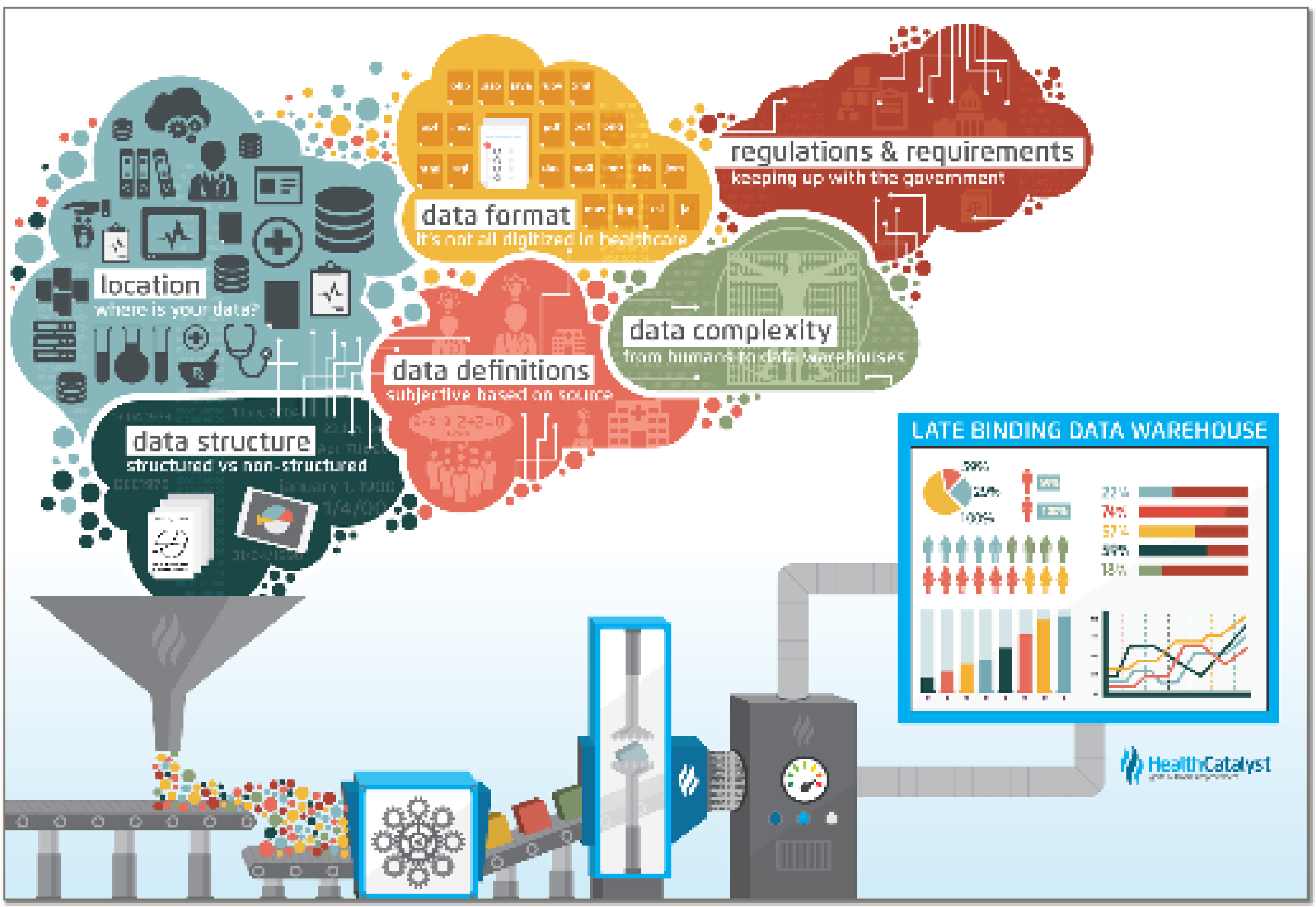
[33] Bates DW et al. Big Data in Health Care: Using Analytics to Identify and Manage High-Risk and High-Cost Patients. *Health Affairs* 2014; 7(2014): 1123-1131.

[34] Feero WG et al. Review Article: Genomic Medicine- An Updated Primer. *N Engl J Med* 2010; 362: 2001-11.

[35] Chan M et al. Smart Wearable Systems: Current Status and Future Challenges. *Artif Intell Med* 2012; 56(3): 137-156.

Healthcare data is unique in several ways that renders the data challenging for AI applications:

1) Location- the data is in various formats (clinical vs. claims data) is often stored in several repositories like clinics, hospitals, and other departments (radiology, laboratories, etc).
2) Data structure- Most (over 80%) of healthcare data, from handwritten doctors' notes to echocardiograms, is unstructured and therefore difficult to handle as a bundle.
3) Data format- Data are in various formats, from digital images to hand-written clinician notes and laboratory reports.
4) Data definitions- data is often defined differently by different clinicians as there is often no universal definition for medical terms.
5) Regulations and requirements- data often are used to fulfill reporting requirements but are often poorly managed.
6) Data complexity- the myriad of data types creates a complex landscape of data that are managed by



various data managers.

An ETL (extract, transform, and load) process is employed in order to extract data out of the system and configure the data for the data warehouse that is favored by business professionals. A data lake is a data storage repository preferred by data scientists and can hold large amounts of raw data, including unstructured data, for later analytic use. There is an important distinction between data warehouse and data lake (see Figure) but perhaps a hybrid “data reservoir” would be the best data repository to take advantage of both types of data storage.

DATA WAREHOUSE	vs.	DATA LAKE
structured, processed	DATA	structured / semi-structured / unstructured, raw
schema-on-write	PROCESSING	schema-on-read
expensive for large data volumes	STORAGE	designed for low-cost storage
less agile, fixed configuration	AGILITY	highly agile, configure and reconfigure as needed
mature	SECURITY	maturing
business professionals	USERS	data scientists et. al.

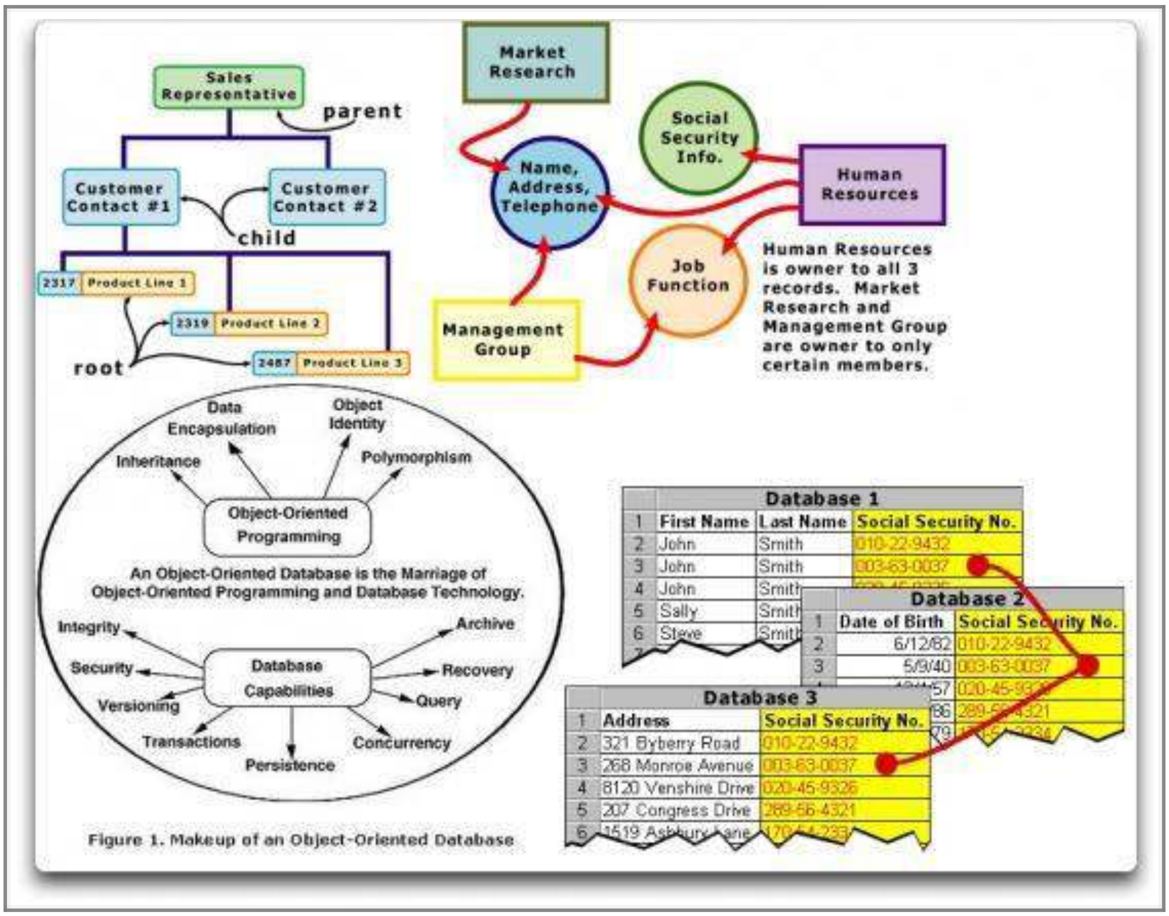
In order to facilitate clinical and administrative data being transferred between software applications, especially with AI-related work, health level-7 (HL7) denotes the seventh (application) level of the International Organization for Standardization (ISO) seven-layer communications model for open systems interconnection (OSI). The HL7 vision is "a world in which everyone can securely access and use the right health data when and where they need it", so AI work in a healthcare organization mandates HL7 as it promotes interoperability of electronic health records. HL7 is sometimes confused with the Healthcare Information and Management Systems Society (HIMSS) Analytics EMR Adoption Model (EMRAM) and its stage designations (see Figure): stage 7 is an environment where paper charts are no longer used (complete EHR) whereas stage 6 is its precursor when healthcare organizations are at the forefront of EHR adoption with interpretable EHR and just prior to stage 7.

EMR Adoption Model℠

Stage	Cumulative Capabilities
Stage 7	Complete EMR; CCD transactions to share data; Data warehousing; Data continuity with ED, ambulatory, OP
Stage 6	Physician documentation (structured templates), full CDSS (variance & compliance), Closed Loop Medication Administration
Stage 5	Full complement of Radiology PACS
Stage 4	CPOE, Clinical Decision Support (clinical protocols)
Stage 3	Nursing/clinical documentation (flow sheets), CDSS (error checking), PACS available outside Radiology
Stage 2	CDR, Controlled Medical Vocabulary, CDS, may have Document Imaging; HIE capable
Stage 1	Ancillaries – Lab, Rad, Pharmacy - All Installed
Stage 0	All Three Ancillaries Not Installed

Healthcare Databases. Most common healthcare database is a relational database and its management system is called a relational database management system (Relational DBMS or simply RDBMS). Oracle and SQL Server are examples of relational DBMS while NoSQL databases such as MongoDB are non-relational DBMS. An online transaction processing (OLTP) database is the predominant use case for RDBMS. A major disadvantage of such a database is that the data is often sequestered. Often, an online analytical processing database (OLAP) is used as an enterprise data warehouse (EDW) solution to solve the problem.

The myriad of database management systems (DBMS) includes (see Figure): hierarchical, network, relational, and object-oriented DBMS. Medical databases have traditionally been primitive flat files with little or no database management and have not advanced far due to electronic medical records only having been recently implemented ([36]). The health care data have therefore been static with indirect sharing mostly through hyperlinks.

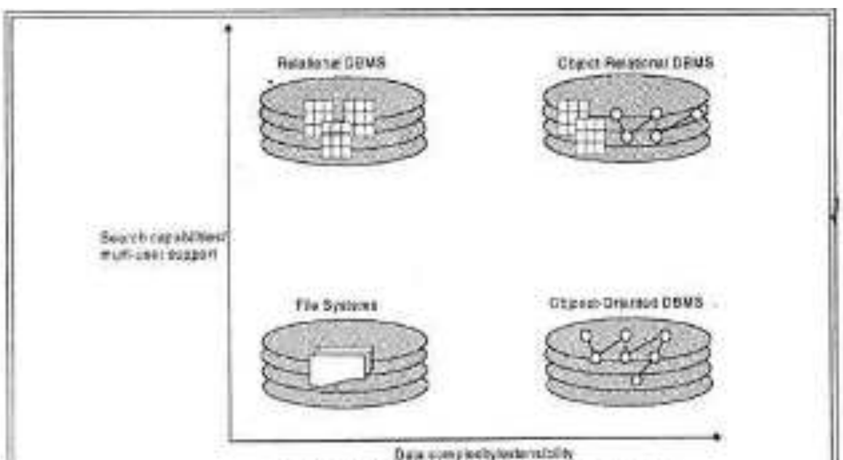


---

36 Mandl KD et al. Escaping the HER Trap- The Future of Health IT. *New Engl J Med* 2012; 366:2240-2242.

In short, most of the present health care data remain embedded in flat files or at best, in relatively simplistic hierarchical or relational DBMS with most of the data centralized and locked into local operating systems that reside in hospitals or offices. There is a paucity of literature on object-oriented approaches in the biomedical area ([37]).

There are limitations to relational DBMS for health care data. Relational DBMS (see figure below) lack sufficient infrastructural support for the larger health care data (such as time-series data, large text documents, and image/videos). In addition, queries are difficult due to the structure of relational DBMS. Structured Query Language (SQL) is used in relational DBMS. Object-oriented DBMS, while more efficient and flexible, lacks the practical functionalities of a relational DBMS especially for search and query functions.



Four-Quadrant View of Database World

A hybrid object-relational DBMS, therefore, can take advantage of strengths from both relational and object-oriented DBMS; it can therefore accommodate the larger, more complex health care data elements while retaining the relational table structure for query purposes (using *Hadoop*, *Oracle*, or *SQL*).
This object-relational DBMS does, however, require more expertise to use, as it is more complex to configure.

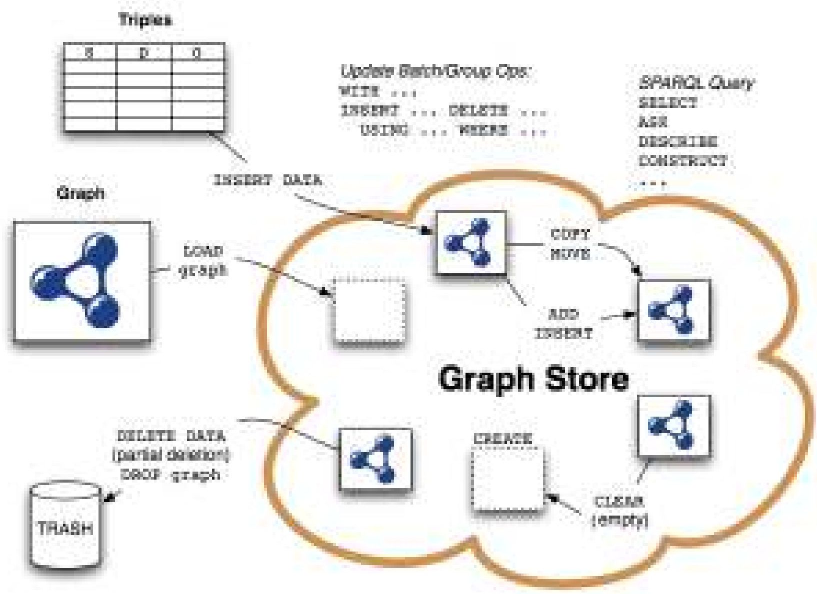
The Not Only SQL (NoSQL) or next generation databases represent databases are characterized by large data volumes, scalable replication and distribution, and efficient queries; these databases are exemplified by document based systems (such as *Mongo DB*) or graph databases and are the future of health care databases.

[37] Gu H et al. Benefits of an Object-Oriented Database Representation for Controlled Medical Terminologies. *J Am Med Inform Assoc* 1999; 6(4): 283-303.

A graph DBMS (used in *LinkedIn* and *Twitter* as well as *Zephyr Health* and *Doximity* and visualized by *Neo4j*) can store data in the form of graph elements (nodes, edges, and properties) in order to facilitate relationship definitions for data elements. This central tenet of delineating connectedness and relationships in a rapid-changing world is often very much needed in biomedicine as quality of care, overall efficiency, and innovation direction become the new paradigm in health care.

In a graph DBMS (see Figure), each data element in the graph will need to be described in the universal language Resource Description Framework (RDF)([38]) as a "triple" (<Subject><Predicate> <Object>). This collection of triples is then stored in a semantic database that can be queried using the semantic version of SQL, SPARQL (Simple Protocol and RDF Query Language).

(Figure: http://www.dajobe.org/talks/201105-sparql-11/sparql-11-graph-store.pn)

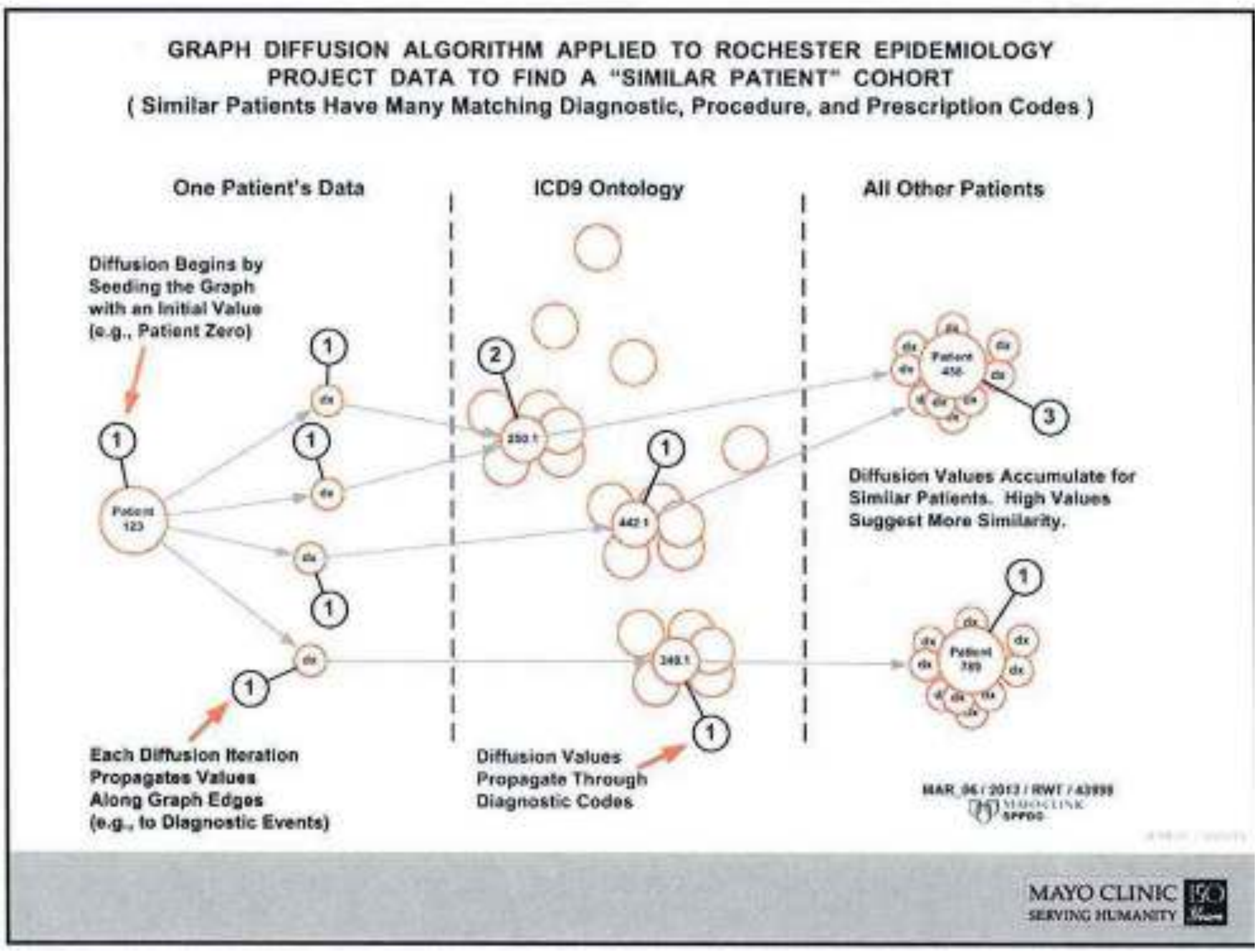


Ontologies and accompanying inference rules can then be embedded in the data to enrich the database.

Algorithms that can be applied to health care data in a graph database include: 1) PageRank- the Google-developed algorithm for placing varying weights on the edges to highlight important nodes; 2) Peer-Pressure Clustering- a graph-based cluster algorithm to find similar groups based on both node and edge data; and 3) Diffusion- an algorithm that can find hidden relationships by defining connectivities in the semantic graph ([39]).

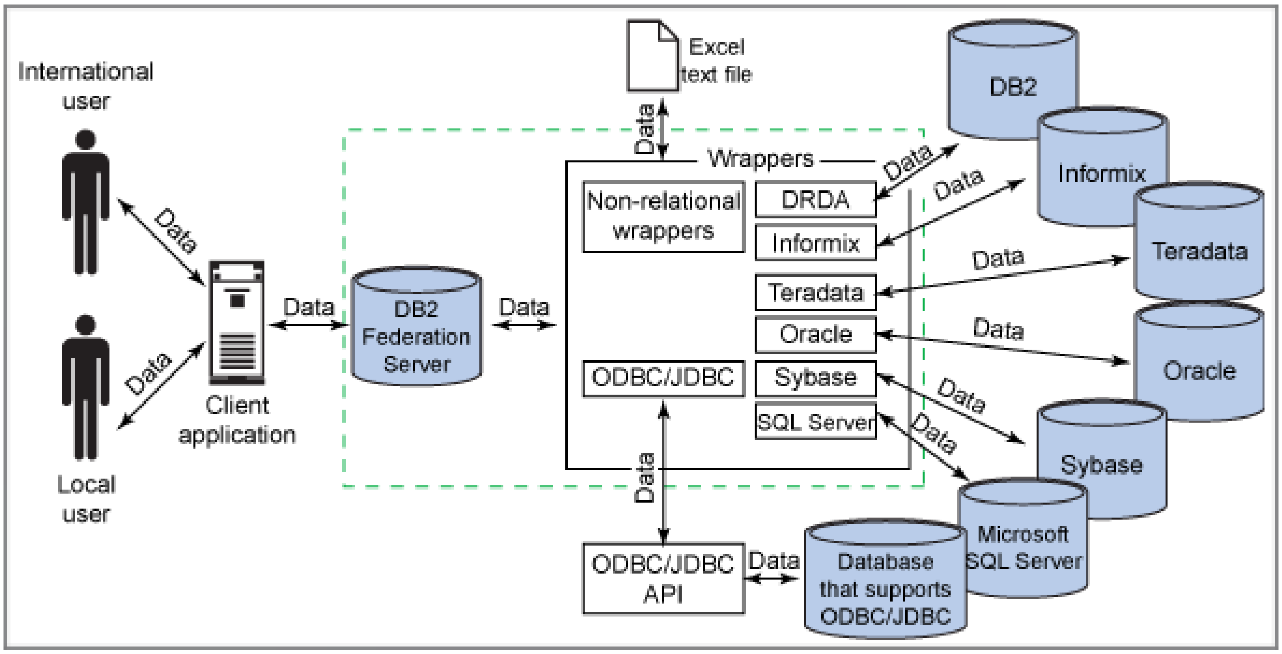
[38] Anguita A et al. Toward a View-Oriented Approach for Aligning RDF-based Biomedical Repositories. *Methods Inf Med* 2014; 53(4) [Epub ahead of print].
[39] Holmes DR. Lecture notes/slides from *The Use of Big Data- Where Are We And What Does the Future Hold?* ISPOR 19th annual meeting, June 2nd, 2014.

The graph DBMS with these search algorithms is especially well designed for complex queries in health care such as chronic disease management, acute epidemiological crises, and health care resource allocation (see Figure). The major limitation of graph DBMS is that it is relatively large and complex, but this limitation can now be partly neutralized with large storage capacity, semantic storage improvements, and superior search algorithms.



The eventual health care databases can be utilization of the above graph DMBS mapped into networks of autonomous databases as a federated or virtual meta-database (or metabase) management system (see figures below for two such systems) for interoperability and collective intelligence ([40])([41]). Many subspecialties lack such a coordinated network of stakeholders and will benefit from such a collaboration.

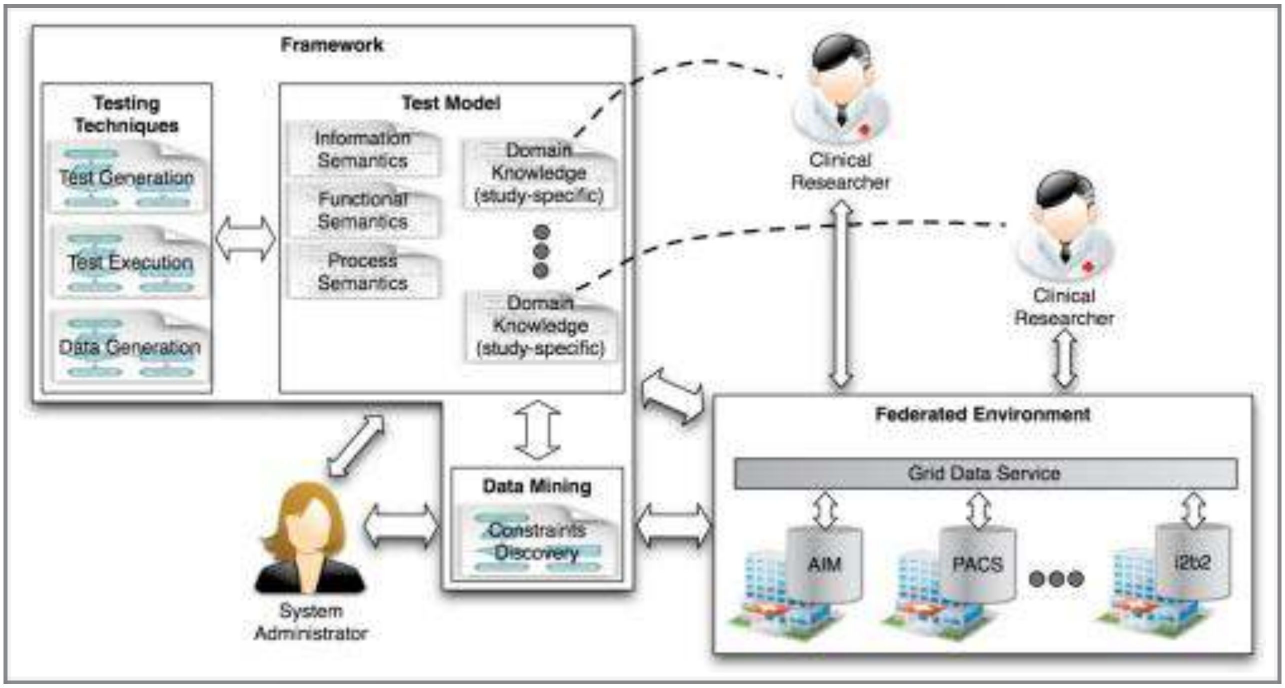
The first figure shows the framework with testing model and technique that are integrated with a data-mining element. The second figure shows the federation server as the center of the distributed heterogeneous data sources architecture. The data is seen to flow from both international as well as local users to



the federation server.

[40] Sinaci AA et al. A Federated Semantic Metadata Registry Framework for Enabling Interoperability Across Clinical Research and Care Domains. *J of Biomed Inform* 2013; 46(2013): 784-794.

[41] Kim M et al. An Informatics Framework for Testing Data Integrity and Correctness of Federated Biomedical Databases. *AMIA Joint Summits on Translational Science Proc* 2011; 2011:22-28.



The advantage of such as network is incorporation of a Web-enabled semantic search and global query capability with data discovery is that it is ideally suited for biomedicine especially with rare diseases and with complex imaging data ([42])([43])([44]). In addition, this federated approach using Internet-based networking technologies can provide excellent collaborative research in epidemiology and public health even at the international level ([45]). Lastly, this federated system provides an excellent framework for the IoT networking paradigm of interconnected smart objects ([46]). This data discovery capability can eventually mature into artificial intelligence possibilities embedded into the database.

[42] Krischer JP et al. The Rare Diseases Clinical Research Network's Organization and Approach to Observational Research and Health Outcomes Research. *J Gen Intern Med* 2014; Suppl 3: 739-744.

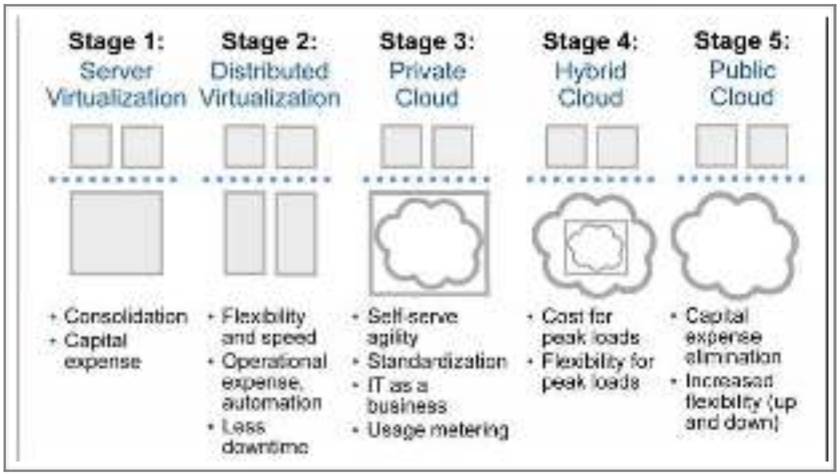
[43] Forrest CB et al. PEDSnet: How a Prototype Pediatric Learning Health System is Being Expanded into a National Network. *Health Affairs* 2014; 7(2014): 1171-1177.

[44] Ozyurt IB et al. Federated Web-Accessible Clinical Data Management Within an Extensible Neuroimaging Database. *Neuroinformatics* 2010; 8(4): 231-249.

[45] Doiron D et al. Data Harmonization and Federated Analysis of Population-Based Studies: the BioSHaRE Project. *Emerging Themes in Epidemiology* 2013; 10:12-20.

[46] Abu-Elkheir M et al. Data Mangement for the Internet of Things: Design Primitives and Solution. *Sensors* 2013; 13(11): 15582-15612.

One key future development of big data analytics is real time analytic processing (RTAP). This is a process in which the data is captured and processed in a streaming fashion using online analytical processing (OLAP) and complex machine learning algorithms ([47]).



Cloud types include public, private, community, and hybrid types. The private cloud (see Figure)(such as hospitals or virtual such as *Verizon*) has advantages that include security and autonomy while the public cloud (such as *Google*) has potential advantages that include scalability and cost-effectiveness.

(Figure: http://cloudblueprint.files.wordpress.com/2011/12/roadmap-to-private-cloud-2.jpg)

Most of biomedicine data is presently stored in local storage or in private clouds due to concern for HIPPA compliance and privacy ([48])([49]). A recent survey of cloud computing adoption in the health care sector revealed that 83% of IT executive in health care reported use of cloud services today with majority utilizing SaaS-based applications (67%)([50]). The cloud infrastructure with convergence to mobile computing, wireless networks, and sensor technology can enable a health care service to be delivered as delineated by Kaur et al in his Cloud Based Intelligent Health Care Service (CBIHCS) for monitoring chronic illnesses such as diabetes ([51]). A community cloud, serving a common interest or purpose, can also be acceptable for a biomedical group or system (such as a specific subspecialty or hospital system)([52]).

---

[47] Branescu I et al. Solutions for Medical Databases Optimal Exploitation. *J Med Life* 2014; 7(1): 109-118.

[48] Regota N et al. Storing and Using Health Data in a Virtual Private Cloud. *J Med Internet Res* 2013; 15(3): e63.

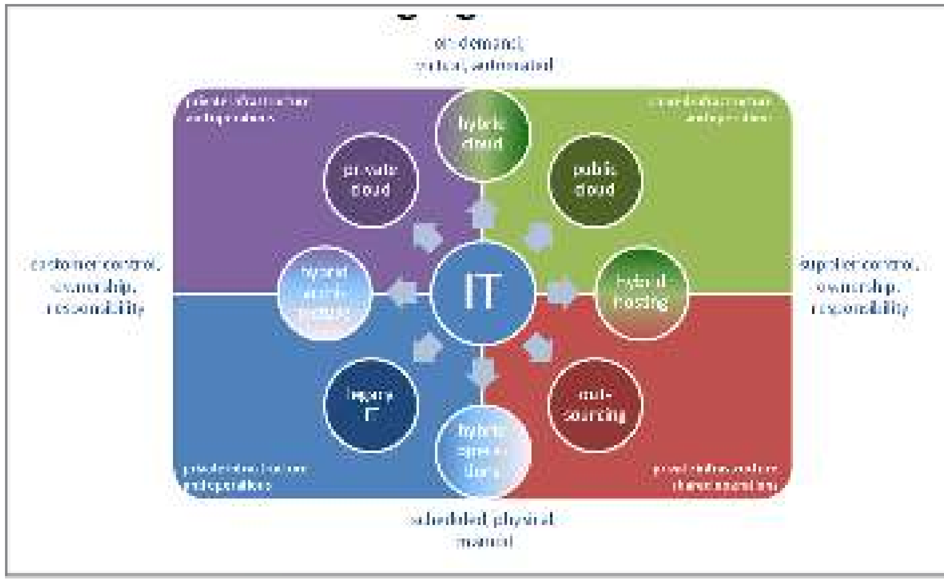
[49] Kaur PD and Chana I. Cloud-Based Intelligent System for Delivering Health Care as a Service. *Comput Methods Programs Biomed* 2014; 113(1): 346-359.

[50] Columbus L. 83% of Healthcare Organizations Are Using Cloud-Based Apps Today. *Technology*, 7/17/2014.

[51] Kaur PD et al. Cloud-based Intelligent System for Delivering Health Care as a Service. Computer Methods and Programs in Biomedicine 2014; 113(2014): 346-359.

[52] Yao Q et al. Cloud-based Hospital Information System as a Service for Grassroots Healthcare Institutions. *J Med Syst* 2014; 38(9): 104-112.

The hybrid cloud has recently emerged as a type of cloud infrastructure that will take advantage of both the public and the private clouds ([53]). The hybrid cloud combines the customization and efficiency as well as security and privacy of the private cloud with the capital preservation and standardization of the public cloud that could be essential for the biomedical milieu ([54]).



(Figure: http://www.bilderbeekconsulting.com/2011/02/four-hybrid-it-environments-it-is-not-just-about-hybrid-cloud)

The usual arrangement is for the private cloud to store data and the public cloud to render services with both communicating via a secured connection. This strategy can also be dynamic with the private cloud utilizing its public cloud partner(s) on an as-needed basis (the so-called "dynamic" hybrid cloud). Logistical challenges of the hybrid cloud include connectivity as well as management of the arrangement on a continuous basis.

The best alternative cloud infrastructure for medical care, however, may be an even more sophisticated cloud infrastructure system that is customized for each need in health care data storage and security (see figure to right). For instance, if supplier control with ownership and responsibility are needed but with shared and private infrastructure, a "hybrid hosting" is desired.

Each configurable cloud infrastructure in the biomedical system should retain the following essential characteristics of cloud computing: on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service ([55]). The cloud infrastructure that is best suited for each situation will be configured to meet that need. Health care with its big data and data analytical needs that concomitantly demands privacy and security can use this flexible cloud infrastructure to meet the challenges.

---

[53] *Your Cloud in Healthcare* by VMware. http://www.vmware.com/files/pdf/VMware-Your-Cloud-in-Healthcare-Industry-Brief.pdf

[54] Nagaty KA. Mobile Health Care on a Secured Hybrid Cloud. *J of Selected Areas in Health Informatics* 2014; 4(2): 1-6.

[55] Barreto D, Lecture for MS&E 238 on July 11, 2014. (Adopted from NIST, 10/09).

Very little in the biomedical data world is truly virtualized at present ([56]). A literature search in PubMed, the health care search engine, yielded no relevant results under "software-defined data center" and "health care". There is, however, a report of software-defined network (SDN) in a hospital setting in Japan using *OpenFlow*™ as the hospital in-house LAN (one of major technical specifications of SDN published by Open Network Foundation)([57]). After the introduction of SDN at this hospital, Nagase avoiding failed route far quicker than conventional technology. This SDN virtualization process simplifies the network environment by two processes: abstraction of network service into policy and automation of application configuration tasks to result in a programmable computer network infrastructure. In short, SDN decouples the control plane from the data plane and creates a more dynamic resource utilization with better oversight of network bandwidth and with direct programmability. One limitation of SDN is its requirement for network architecture and design skills.

In addition, the storage of heterogeneous health care data sources can be abstracted into designated data pools, with this process being an app-centric policy-based automation (software-defined storage)(such as *IBM*'s *Big Blue Elastic Storage*). This strategy is ideal for the biomedical system with its data-intensive applications that require immediate access and rapid analytics as well as matching the storage capability with the type of data ([58]). As an example, this strategy has been implemented at Maimonides Medical Center with *DataCore's SANsymphone-V* software with savings in hardware and personnel costs and increased performance of applications ([59]).

---

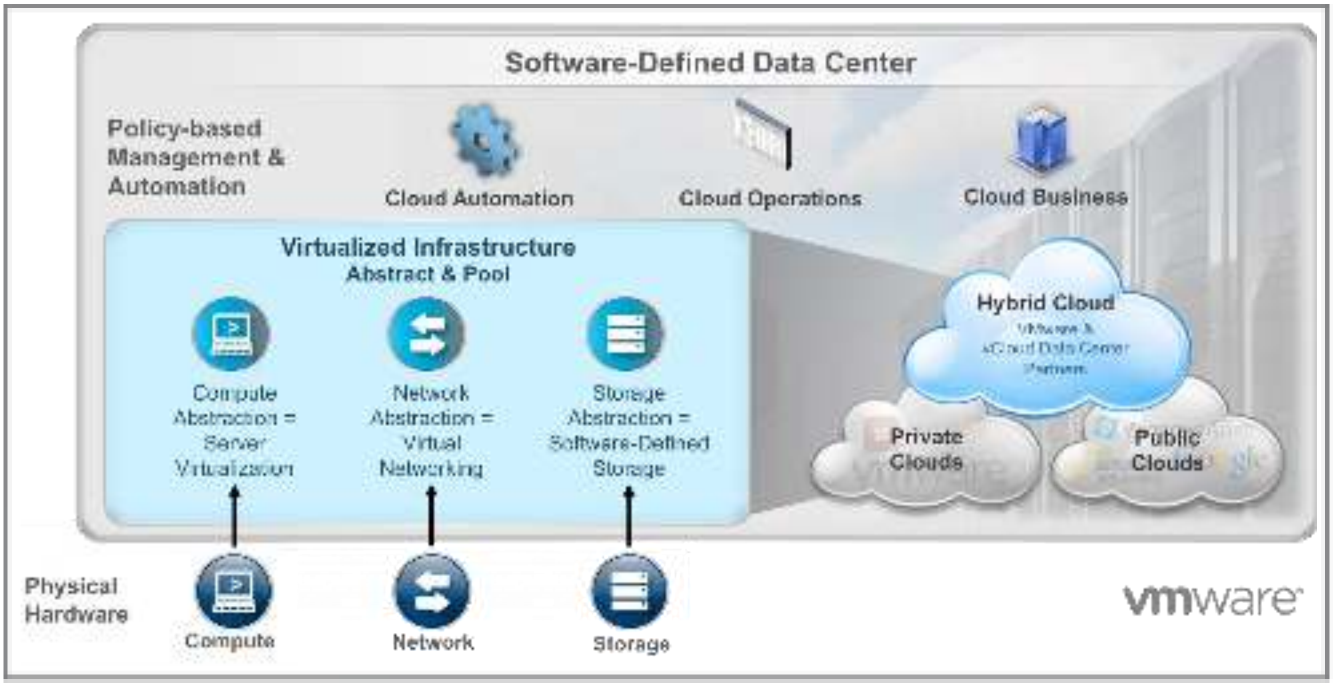
[56] Personal communication with Dr. Spyro Mousses, July 28th, 2014.

[57] Nagase K. Software Defined Network Application in Hospital. *J of Innovation Impact* 2013; 6(1): 1-11.

[58] Personal communication with Dr. Marty Kohn (formerly of IBM), July 9th, 2014.

[59] How Software-Defined Storage Brought Maimonides Medical Center to the Forefront of Healthcare IT. *The Virtual Viewpoint*, July 30, 2104.

The aforementioned SDN concept has evolved into a software-defined data center (SDDC) architecture (with server virtualization, software-defined networking, and storage hypervisor) that can be entirely virtualized so that all of the infrastructure and the management can be automated by software. In short,



the SDDC frees the application layer from the hardware layer (see Figure).

An entirely virtualized infrastructure can include compute, network, security, and storage abstractions such that it is IT as a service (ITaaS) and its cloud infrastructure also managed by automation.

One advantage of such a system is its remote programmability to render it agile and automated as well as global and continuous. All of the components are essentially decoupled from the hardware as a ubiquitous software for all users in various hospitals and programs at any given time. Another advantage is the accelerated service delivery (from weeks/days to even hours). The dynamic configuration of SDDC can optimize resource allocation and improve efficiency in health care.

The future of software-define data systems will be a federated system with more standardized network protocols and more automated interfaces for management. With emerging sophisticated database management systems and cloud and virtual computing technology, medical data can be efficiently organized into a virtual intelligent biomedical data "ecosystem" to better serve the needs of hospitals and health systems ([60]). In addition, the wireless sensor networks as well as other patient-generated data will need to be virtualized to provide effective solutions in health care ([61])([62]). This virtual strategy will enable medical data to converge with artificial intelligence methodologies to promulgate true medical intelligence in the cloud ([63]).

At the virtualization stage, several artificial intelligence methodologies can be implemented to form a collective intelligence ensemble:
One future data analytic tool is predictive modeling for real-time clinical decision making called electronic health care predictive analytics (or e-HPA) to fulfill the triple aim: improving outcomes, enhancing patient experience, and reducing health care costs ([64]). Another essential arm of health care intelligence will be data visualization ([65]). Graph-based data analysis with real-time structured network analysis (SNA) is an excellent modeling of relationships and is highly effective. Natural language processing (NLP) will also be a valuable part of the portfolio of artificial intelligence methodologies in interpreting data ([66]). Lastly, one of the most promising techniques is deep learning in the form of deep belief networks that consist of several layers of neural networks ([67]). The recent advent of TrueNorth, the neuromorphic IBM chip can accelerate the quest for artificial intelligence.

---

[60] Graschew G et al. New Trends in the Virtualization of Hospitals- Tools for Global e-Health. *Stud Health Technol Inform* 2006; 121: 168-175.

[61] Islam MM et al. A Survey on Virtualization of Wireless Sensor Networks. *Sensors* 2012; 12(2): 2175-2207.

[62] Howie L et al. Assessing the Value of Patient-Generated Data to Comparative Effective Research. *Health Affairs* 2014; 7(2014): 1220-1228.

[63] Scott DJ et al. Accessing the Public MIMIC-II Intensive Care Relational Database for Clinical Research. *BMC Med Inform Decis Mak* 2013; 13:9.

[64] Amarasingham R et al. Implementing Electronic Health Care Predictive Analytics: Considerations and Challenges. *Health Affairs* 2014; 7(2014): 1148-1154.

[65] Data Visualization in Burlingame N et al(eds) *A Simple Introduction to Data Science.* New Street Communications, LLC, Wickford, RI, 2012.

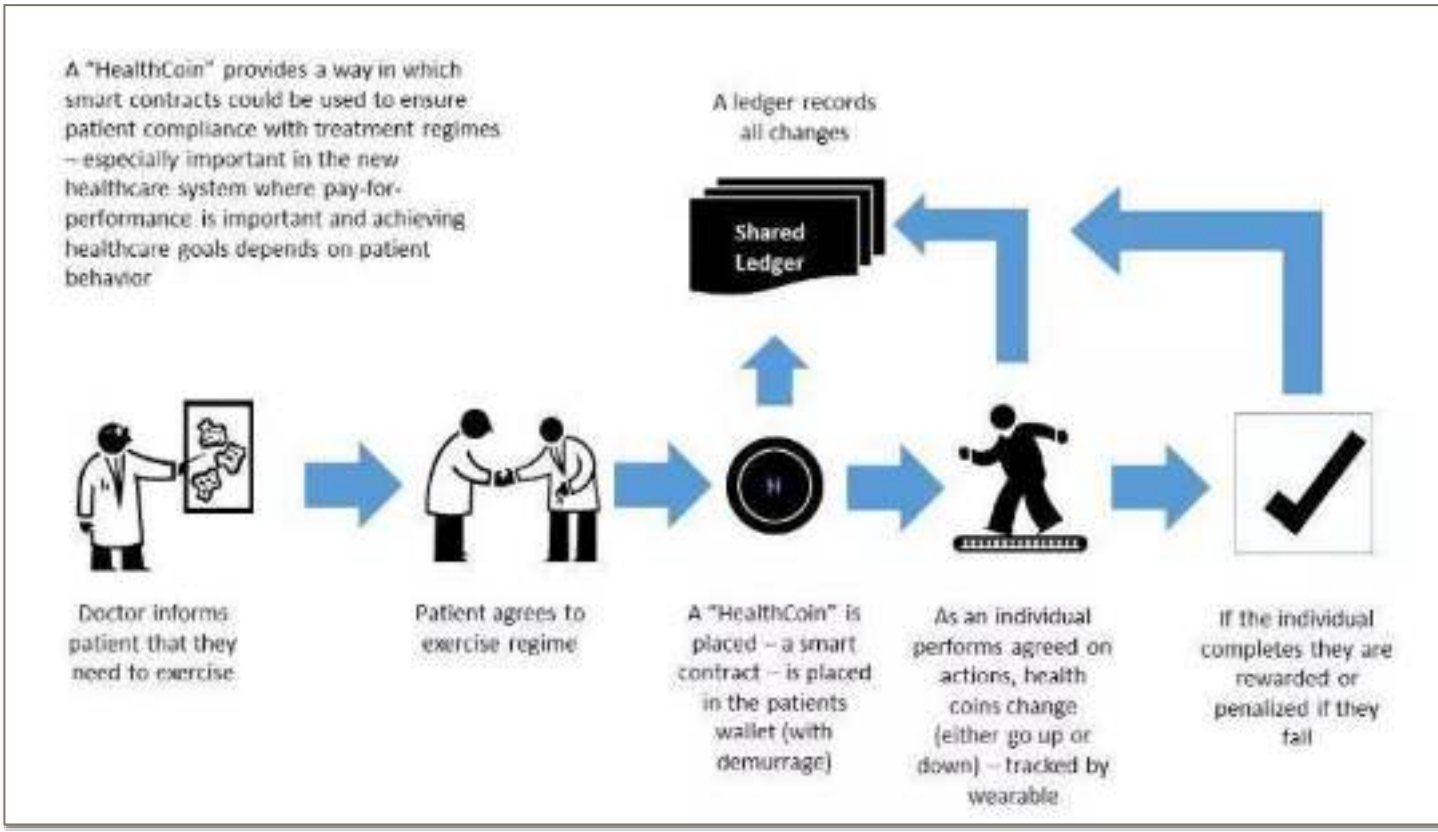
[66] Doan S et al. Natural Language Processing in Biomedicine: A Unified System Architecture Overview. *Methods Mol Biol* 2014; 1168: 275-294.

[67] Arel I et al. Deep Machine Learning- A New Frontier in Artificial Intelligence Research. *IEEE Computational Intelligence* 2010; 13-18.

One major area of concern will always be data security in the cloud as present data shows 94% of health care institutions having had breach of data ([68]). Additional future cloud and data security concepts to be adopted for the future will need to include novel concepts as blockchain (see Figure) as well as:

1) Homomorphic encryption. This is an encryption strategy that allows for certain computations to be performed on medical data while they are still encrypted ([69]). One significant limitation for this security solution is that the processing speed is slowed during this process.

2) Differential privacy. This security process uses sophisticated algorithms to add sufficient "noise" to the



data to render it less vulnerable to linkages to other databases for matching purposes.

---

[68] Perakslis ED. Cybersecurity in Health Care. *N Engl J Med* 2014; 371(5): 395-397.

[69] Kocaba O et al. Medical Data Analytics in the Cloud Using Homomorphic Encryption in Chelliah PR et al (eds) *Handbook of Research on Cloud Infrastructures for Big Data Analytics*. ITI Global.

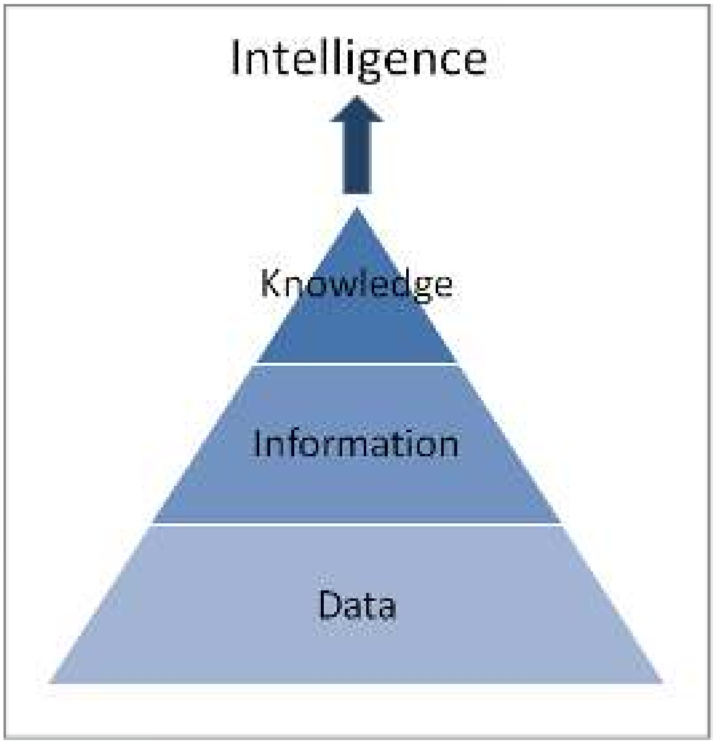
The exponential convergence of existing biomedical data with both genomic and biophysiologic data will render medical data to be even more voluminous, complex, and heterogeneous. This explosion of medical data will need a more sophisticated database management strategy as well as cloud and virtual environments to enhance data discovery as well as ensure data security and privacy.

First, future medical data can be managed in a graph-based meta-database management system with real time analytic processing for both its storage capability and its query flexibility to accommodate the large and complex medical data in the ensuing decades.

Second, future medical data in the customized cloud infrastructure will be far more sophisticated than a simple public-private dichotomy and can be customized from a cloud infrastructure system based on customer vs supplier control, ownership, and responsibility as well as private vs shared infrastructure and operations; cloud security can be further enhanced by mechanisms such as homomorphic encryption and differential privacy.

Third, a software-defined data center (SDDC) architecture can be entirely virtualized so that the infrastructure that includes compute, network, and storage abstractions will result in IT as a service (ITaaS). The future medical data system will be entirely in a virtual synergy with humans and contribute to medical intelligence.

The future of biomedicine can include a proposal for an artificial intelligence-inspired cloud continuum of data-information-knowledge-intelligence (a “medical intelligence” as a service, or “MIaaS”). By embedding intelligence into all aspects of medical data from graph database and meta-database management system to customized cloud infrastructure and to software-defined data center and virtualization, the aforementioned strategies can accelerate this transformation in biomedicine from fragmented and unstructured data sets to cohesive and agile information imbued with medical intelligence.



The Data-to-Intelligence Pyramid. The continuum from data to information and from knowledge to intelligence starts with precise database management and accurate data analysis. To have good AI or intelligence in healthcare, one must start with good healthcare data (the foundation layer of the data-intelligence pyramid)(see Figure). Data is processed and interpreted to induce meaning, and this leads to the next level, information. While computers need data, humans need information.

From information, one attains knowledge, which is derived from one's experience and analysis. Intelligence is therefore the ability as well as the velocity to apply this knowledge. Wisdom (not pictured) is thought to be a quality of "knowing" without necessarily having the logic to affirm the observation or decision.

Present day AI methodologies, especially deep and reinforcement learning (see below), are changing the "human" role and the machine expectations in this data-to-intelligence pyramid.

Deep Learning and Healthcare

The application of deep learning in healthcare and medicine has exponentially escalated in the past few years. The number of publications have increased from only 122 even in 2007 to now well over 1,000 in 2017, and the increase is probably the most dramatic in radiology (4 in 2007 to over 100 in 2017).

Healthcare and medical applications of deep learning is in its nascent stages. Deep learning will be increasingly applied to escalating data in healthcare and medicine as these layers can be assigned to the many genomic (such as DNA sequence, regulatory features, etc)([70]) as well as phenotypic expressions (such as clinical measurements, biomarkers, imaging data, disease subtypes)([71]). A prime example of the use of deep learning and clinician-data scientist collaboration is the recent National Health Services (NHS) and Google DeepMind collaboration for eye disease detection in over a million patients (see next section on Applications). Another example is the skin cancer work by the Stanford group that was published in *Nature* (see below). Finally, Dudley's group elegantly described the "deep patient" concept in which a patient's electronic medical records is used to predict the patient's future disease profile ([72]).

Overall, the fastest growing trends in deep learning and medicine and healthcare are observed in the following areas:
1) Medical Imaging: Use of GPU-accelerated deep learning and computer vision with classification, detection, and segmentation can automate analysis of a myriad of medical images such as CT, MRI, X-rays, and even moving images such as echocardiograms and angiograms ([73]).
2) Decision Support: Deep learning techniques can be used to collate and analyze the heterogeneous electronic health record data pools such as doctors notes, laboratory data, drug information, and medical images in order to facilitate diagnosis and therapy. A point-of-care capability can then be devised.
3) Precision Medicine: The genomic sequencing data and phenotypic expression information can be coupled to devise an individualized diagnosis and therapy for patients; population health can also benefit from a generalized approach.
Deep learning can also be applied to other domains such as drug discovery and event prediction.

[70] Rusk N. Deep Learning. *Nature Methods* 2016; 13 (1): 35.

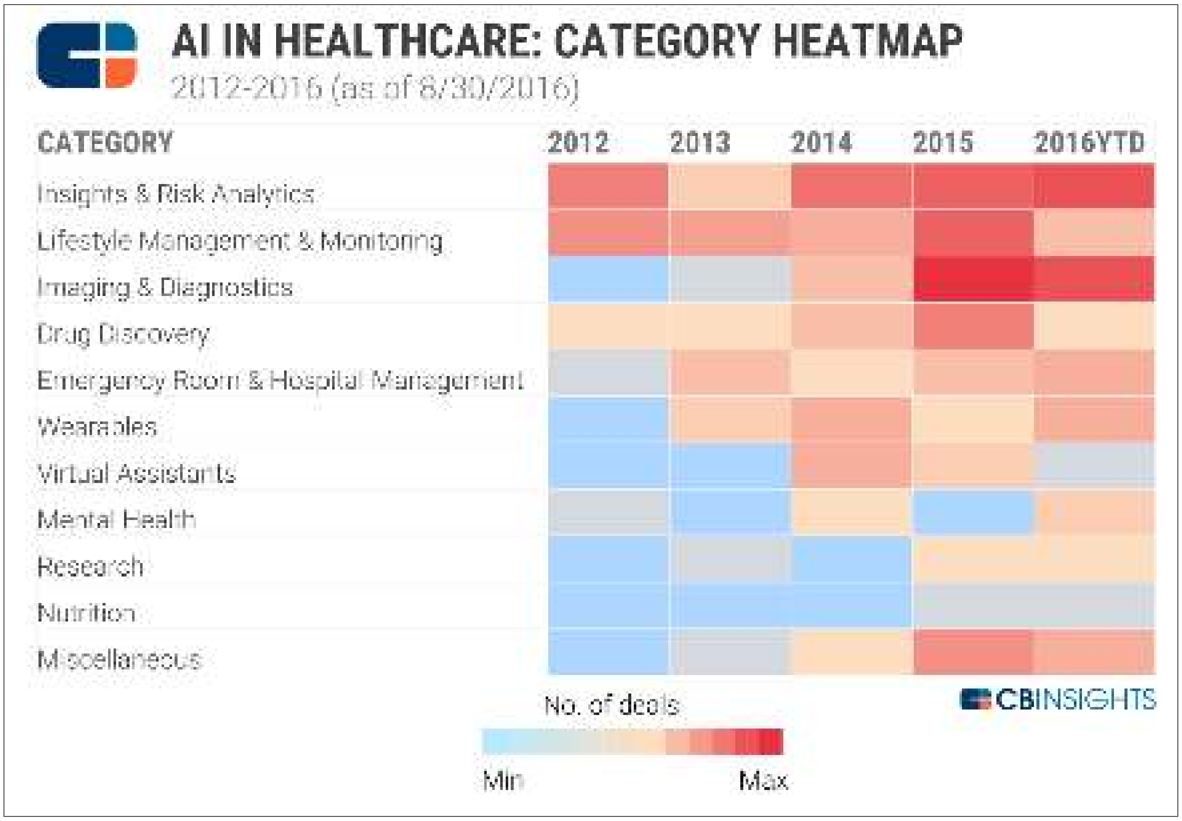
[71] Deo RC. Machine Learning in Medicine. *Circulation* 2015; 132:1920-1930.

[72] Miotto R, Li L, Kidd BA et al. Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records. *Nature Scientific Reports* 2016; 6:26094 (1-10).

[73] Greenspan H, van Ginneken B, and Summers RM. Guest Editorial/ Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Transactions on Medical Imaging* 2016; 35(5): 1153-1159.

Current Applications of Artificial Intelligence in Medicine

There has been more available resources for AI in medicine and healthcare than ever before. In addition to the proliferation of AI tools and availability of educated personnel, venture capital has also escalated into high levels and is now about 15% of all AI-related deals (see Figure).



Of 12 categories of AI in healthcare, the most active sector in terms of number of deals in 2016 is imaging and diagnostics (with about one-third of all deals), followed by insights and risk analytics (see Figure). Other areas that are active include: lifestyle management and monitoring; emergency room and hospital management; wearables; and mental health projects.

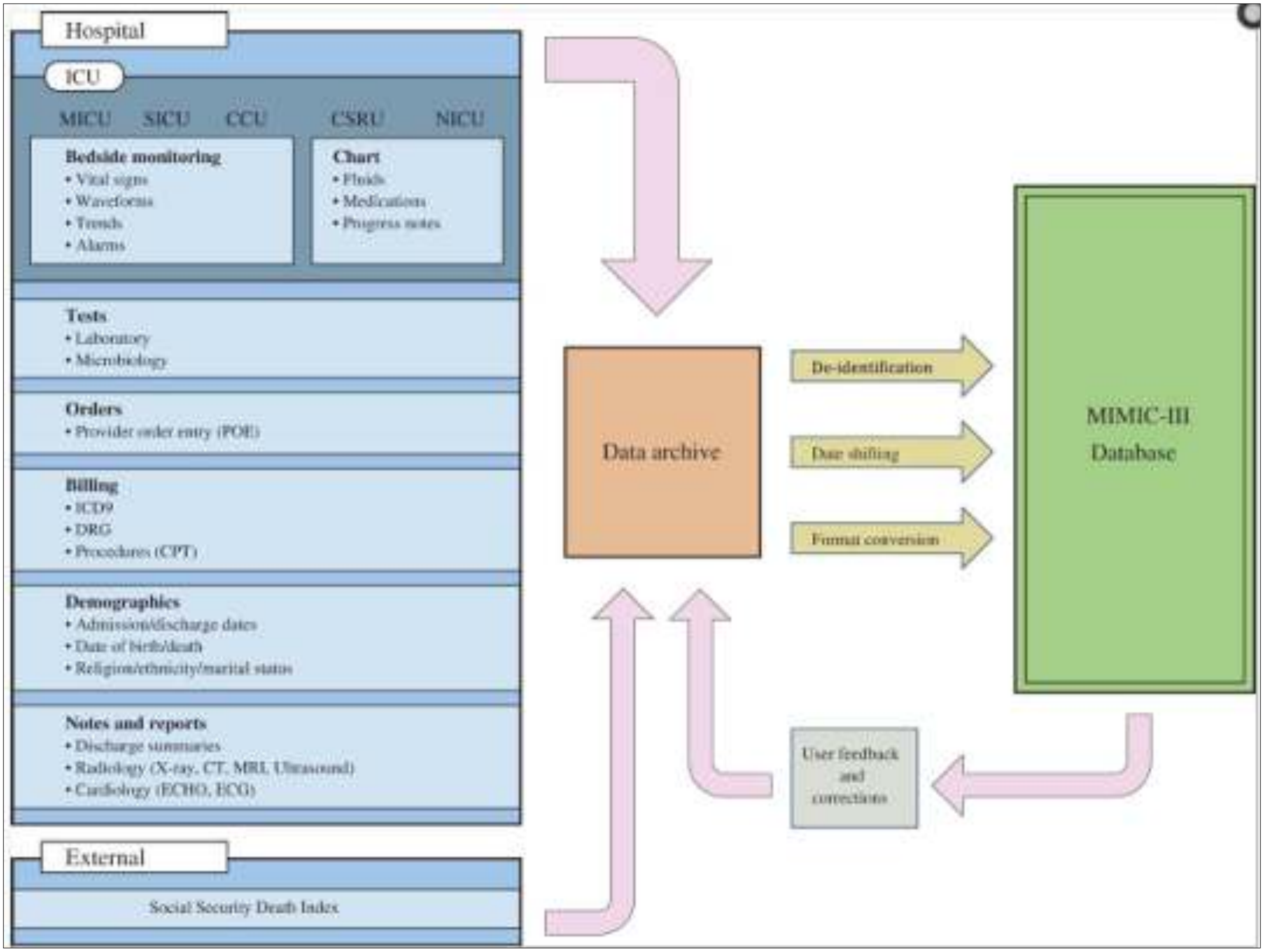
The present state of artificial intelligence in medicine includes a myriad of applications:

Decision Support and Hospital Monitoring. A recent review of the history of clinical decision support (CDS) states the dramatic improvement in this sector due to the advent of cognitive aids to support diagnosis, treatment, care-coordination, surveillance and prevention, and health maintenance or wellness ([74]).

An example of application of artificial intelligence in the hospital setting is the use of machine learning algorithms for training expert-labeled vital sign data streams to automatically classify vital sign alerts as real or artifact in order to clean such data for future modeling ([75]).

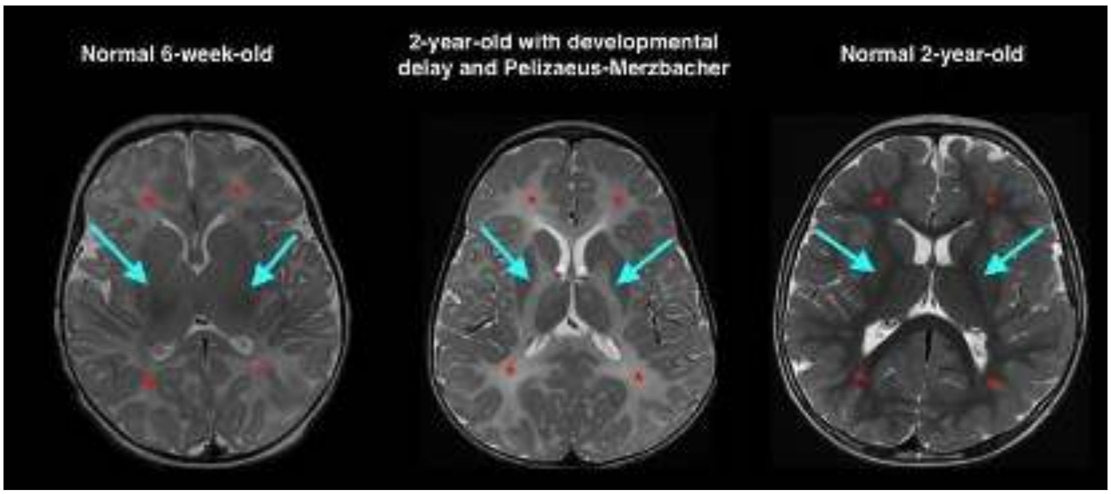
Another robust illustration of the value of databases coupled with analytics is the Medical Information Mart for Intensive Care (MIMIC) accessible critical care database ([76])(see Figure). MIMIC is a large, single-center database comprising of information in the critical care unit from admitted patients for data mining and modeling of conditions that resulted in many publications. This publicly available database has been even used for datathons around the world by the MIT group to lead multidisciplinary groups that consist of clinicians and data scientists.

[74] Middleton B, Sittig DF, and Wright A. Clinical Decision Support: A 25 Year Retrospective and a 25 Year Vision. *Yearbook Med Inform* 2016; Suppl 1: S103-116.
[75] Hravnak M, Chen L, Dubrawski A et al. Real Alerts and Artifact Classification in Archived Multi-Signal Vital Sign Monitoring Data: Implications for Mining Big Data. *J Clin Monit Comput* 2016; 30(6): 875-888.
[76] Johnson AE, Pollard TJ, Shen L et al. MIMIC-III, A Freely Accessible Critical Care Database. *Sci Data* 2016; 3:160036.


Hospital
ICU
MICU
SICU
CCU
CSRU
NICU
Bedside monitoring
• Vital signs
• Waveforms
• Trends
• Alarms
Chart
• Fluids
• Medications
• Progress notes
Tests
• Laboratory
• Microbiology
Orders
• Provider order entry (POE)
Billing
• ICD9
• DRG
• Procedures (CPT)
Demographics
• Admission/discharge dates
• Date of birth/death
• Religion/ethnicity/marital status
Notes and reports
• Discharge summaries
• Radiology (X-ray, CT, MRI, Ultrasound)
• Cardiology (ECHO, ECG)
External
Social Security Death Index
Data archive
De-identification
Date shifting
Format conversion
MIMIC-III
Database
User feedback
and
corrections

Medical Imaging and Biomedical Diagnostics. There is much promise in the utilization of AI methodologies such as deep learning for automated and/or augmented biomedical image interpretation in radiology, pathology, dermatology, ophthalmology and cardiology in the near future.

AI techniques such as artificial neural networks, support vector machines, classification tree and ensemble methods such as random forest are able to be applied to molecular imaging modalities in clinical diseases such as neurodegenerative diseases ([77]). This study demonstrated that computer aided

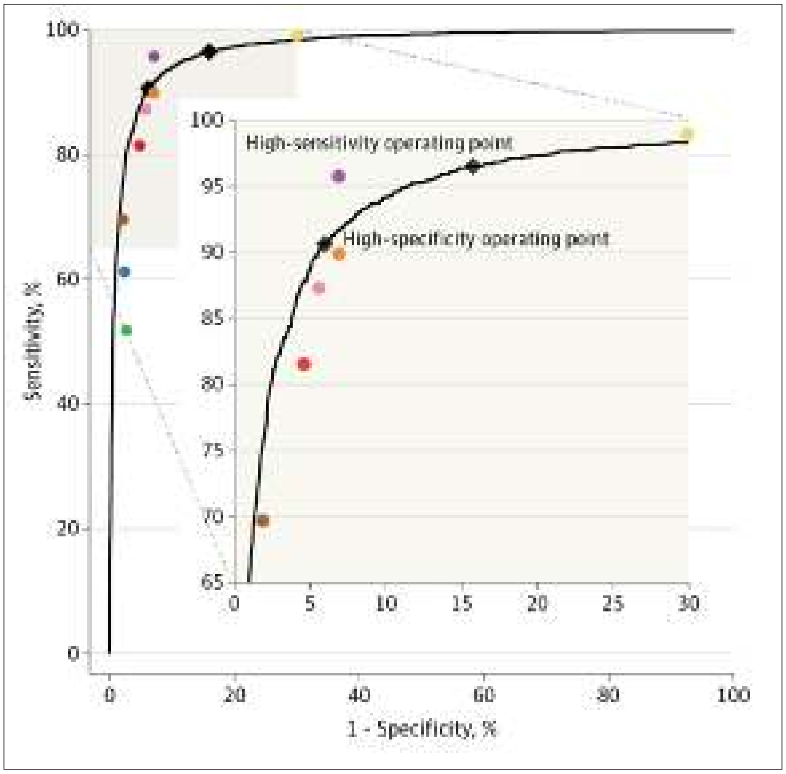


diagnosis systems such as this are very promising to contribute to the diagnostic process.

In addition, GE Healthcare has very recently partnered with Boston Children's Hospital to develop an AI-backed decision support platform with cloud computing for the diagnosis and treatment of pediatric brain diseases ([78]). This AI strategy of having a reference library of normal pediatric MRI brain scans coupled with deep learning is particularly helpful for pediatric brain imaging as disorders can be misinterpreted as normal brain maturation while the opposite is also true (normal brain maturation can be misdiagnosed as abnormal changes and lead to misdiagnoses and inappropriate treatment)(see Figure). This decision support platform will be available worldwide for all pediatric brain imaging as pediatric neuroradiologists are scarce.

[77] Cascianelli S, Scialpi M, Amici S et al. Role of Artificial Intelligence Techniques (Automatic Classifiers) in Molecular Imaging Modalities in Neurodegenerative Diseases. *Curr Alzheimer Res* 2016; Jun 20 [Epub ahead of print].

[78] Al Idrus A. GE, Boston Children's to Create Deep Learning Tool for Pediatric Brain Scans. *FierceBiotech,* November 28, 2016.



Even more newsworthy is the November 2016 publication in the *Journal of the American Medical Association* on algorithms based on deep convolutional neural network devised by Google to detect diabetic retinopathy in millions of retinal fundus photographs ([79]). The study showed that the algorithm is as good as board-certified ophthalmologists in making the diagnosis with high specificity and sensitivity and an area under the receiver operating curve of 0.99 (see Figure). The performance of the algorithm for automated diagnoses (black curve with black diamonds) is compared with 8 ophthalmologists with manual grading (colored circles).

In biomedical diagnostics, medical geneticists are often frustrated by the tedious nature of genotype-phenotype interrelationships among syndromes, especially for extremely rare syndromes. Now, medical geneticists are able to use a visual diagnostic decision support system that employs machine learning algorithms and digital imaging processing techniques in a hybrid approach for automated diagnosis in medical genetics ([80]).

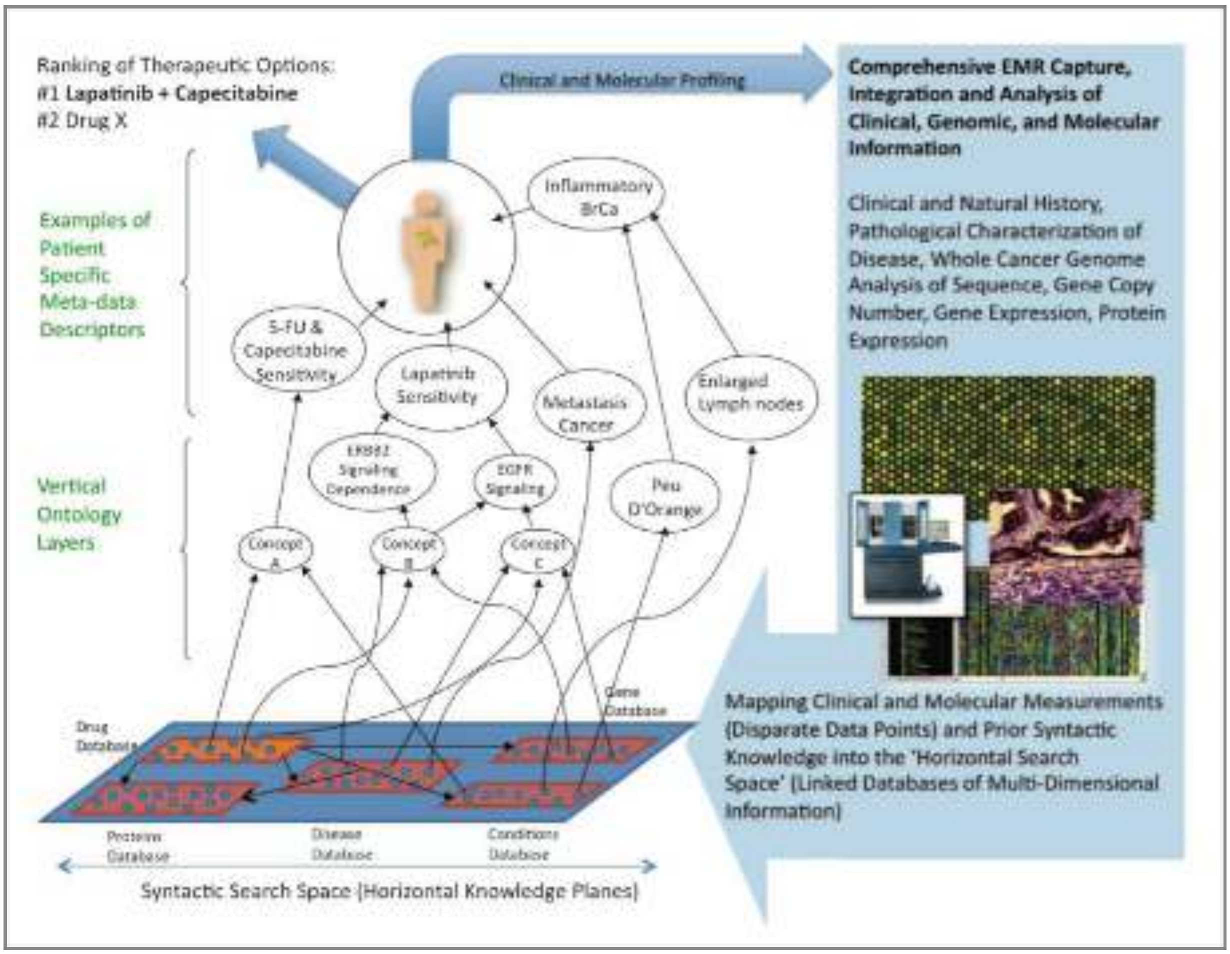
---

[79] Gulshan V, Peng L, Coram M et al. Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. *JAMA* 2016. do:10.1001/jama.2016.17216.

[80] Kuru K, Niranjan M, Tunca Y et al. Biomedical Visual Data Analysis to Build an Intelligent Diagnostic Decision Support System in Medical Genetics. *Artif Intell Med* 2014; 62(2): 105-118.

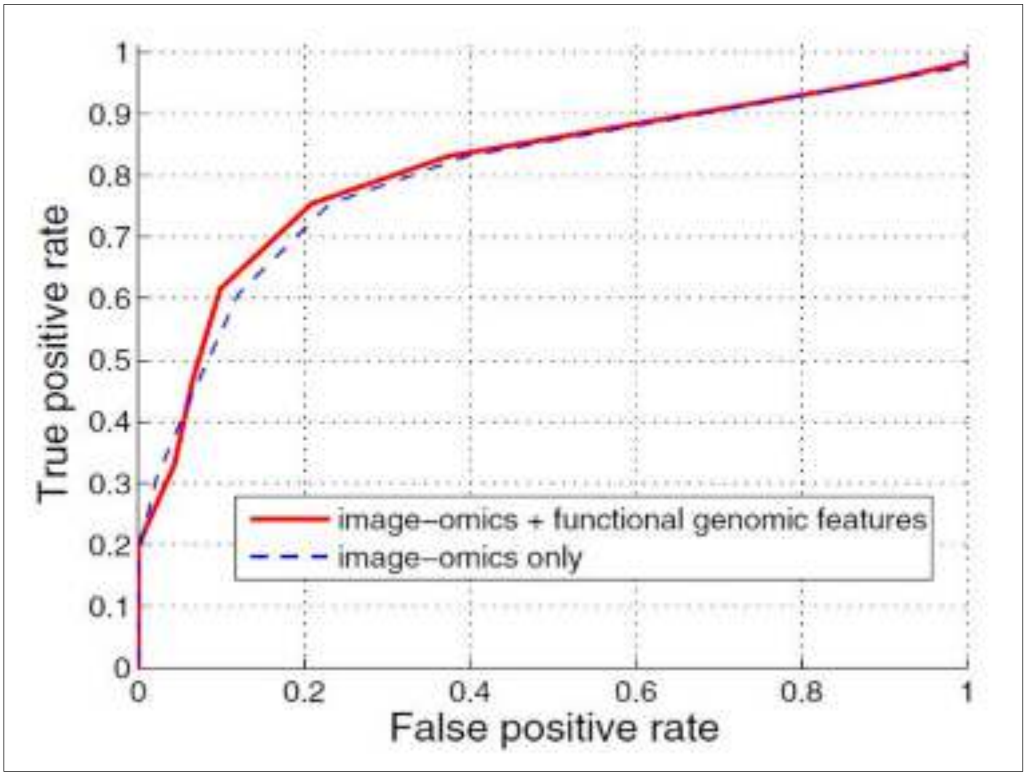
Precision Medicine and Drug Discovery. Precision medicine with its complexity and enormity of data to be analyzed is particularly well suited for the portfolio of AI methodologies such as deep learning as similar patients can be identified and assessed. Precision medicine at its highest level will need a disruptive computational platform for new biomedical knowledge discovery.

One such proposal is the BioIntelligence Framework proposed by Farley et al ([81]). In this model, a scalable computational framework leverages a hypergraph-based data model and query language that may be suited for representing complex multi-lateral, multi-scalar, and multi-dimensional relationships (see Figure). This hypergraph-like store of public knowledge is coupled with an individual's genomic and other patient information (such as imaging data) to drive a personalized genome-based knowledge store for clinical translation and discovery. Patients of very similar genomic and clinical elements can be



discovered and matched for diagnostic and therapeutic strategies (see Figure).

[81] Farley T, Kiefer J, Lee P et al. The BioIntelligence Framework: A New Computational Platform for Biomedical Knowledge Computing. *J Am Med Inform Assoc* 2013; 20(1): 128-133.

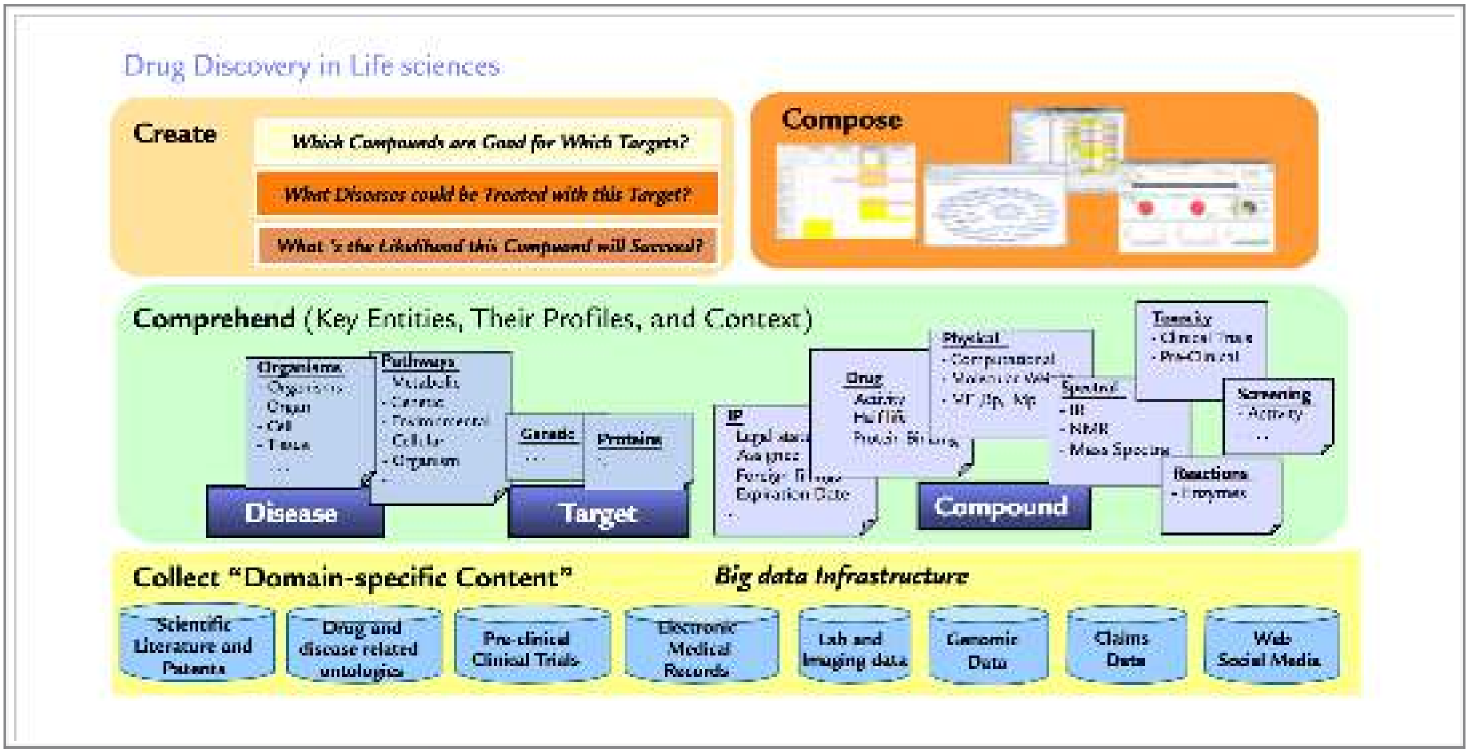


In breast cancer, the combination of computer-aided diagnosis of image-omics (pathological images) and functional genomic features improved the classification accuracy by 3% ([82])(see Figure).
In this study, support vector machine for differentiating stage I breast cancer from other stages are learned with use of computer-aided diagnosis that enables joint analysis of functional genomic information and image from pathological images. The entire biomedical imaging informatics framework consisted of image extraction, feature combination, and classification.

In addition, there are many potential applications of deep learning for large datasets in pharmaceutical research (such as physicochemical property prediction, formulation prediction, and properties such as absorption, distribution, metabolism, excretion, toxicity, and even target prediction)([83]). An essential part of the precision medicine paradigm is individualized therapy based on genotype-phenotype coupling and pharmacogenomic profiles.

[82] Su H, Shen Y, Xing F et al. Robust Automatic Breast Cancer Staging Using a Combination of Functional Genomics and Image-omics. *Conf Proc IEEE Eng Med Biol Soc* 2015; 2015: 7226-9.
[83] Ekins S. The Next Era: Deep Learning in Pharmaceutical Research. *Pharm Res* 2016; 33(11): 2594-2603.

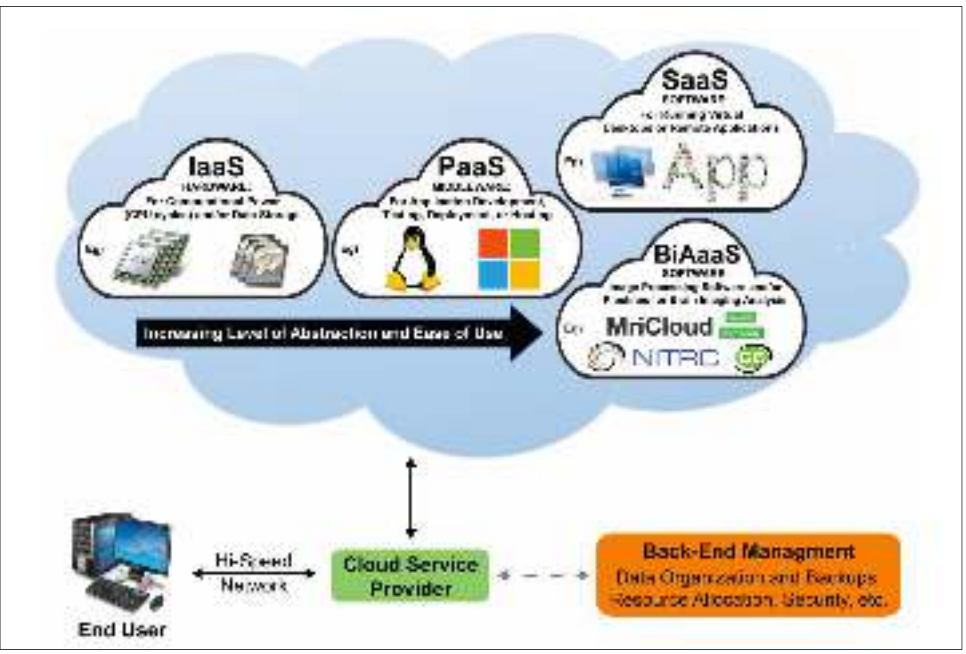
Finally, cognitive solutions are designed to fully integrate and analyze relatively large data sets such as in life sciences for drug discovery (see Figure). Such strategies include collecting domain specific content in the form of scientific literature and patents, drug and disease related ontologies, pre-clinical clinical trials, electronic medical records, labs and imaging data, genomic data, and even claims data and social media data.



Cloud Computing and Big Data. Cloud computing in healthcare is at present in the form of singular, individual features such as elasticity, pay-per-use and broad network access rather than as cloud paradigm on its own ([84]). Cloud computing is therefore often in the "OMICS-context" with computing in genomics, proteomics, and molecular medicine with little use in other domains.

Although the storage of genomic data into the public cloud raises issues such as form of security and privacy, initial efforts to secure computation techniques that can enable comparative analysis of human genomes have been productive. One such effort is the NIH-funded National Center for Biomedical Computing iDASH (integrating Data for Analysis, anonymization, and Sharing) and its Critical Assessment of Data Privacy and Protection competition to evaluate the capacity of the cryptographic technologies for protecting computation over human genomes in the cloud while promoting cross-institutional collaboration ([85]).

In addition, high-performance computing platforms such as clusters, grids and clouds can be used by neurologists, radiologists, and researchers for imaging such as neuroimaging to increase both storage and/or computational power ([86])(see Figure).



Finally, IBM Watson debuted SleepHealth, an app and ResearchKit (Apple) study designed to investigate the possible connection between sleep habits and health. This app requires the support of HIPPA-compliant Watson Health Cloud and will gather the crowd-sourcing data for an unprecedented study of sleep and health.

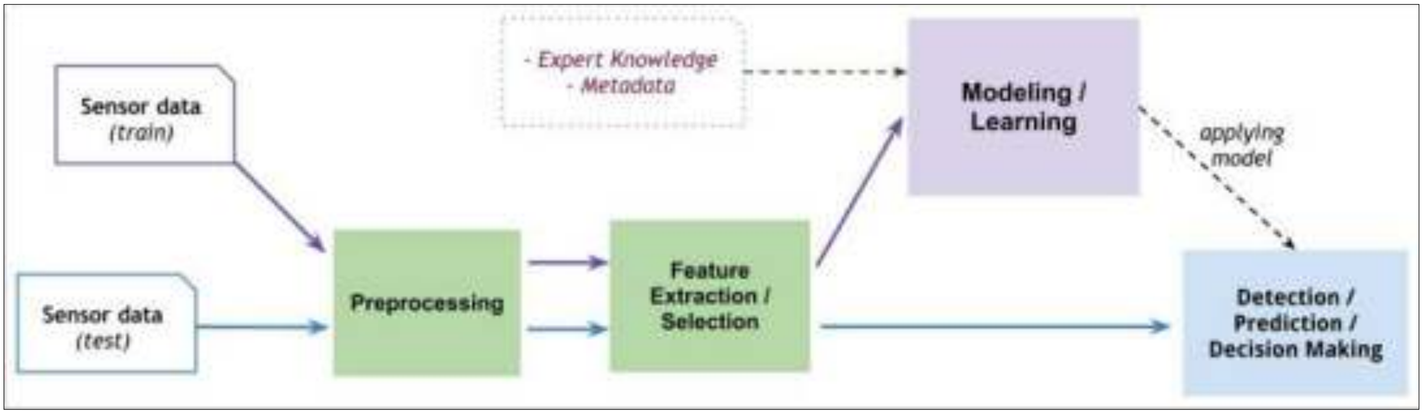
---

84 Griebel L, Prokosch HU, Kopcke F et al. A Scoping Review of Cloud Computing in Healthcare. *BMC Med Inform Decis Mak* 2015; 15:17.

85 Tang H, Jiang X, Wang X et al. Protecting Genomic Data Analytics in the Cloud: State of the Art and Opportunities. *BMC Med Genomics* 2016; 9(1): 63.

86 Shatil AS, Younas S, Pourreza H et al. Heads in the Cloud: A Primer on Neuroimaging Applications of Performance Computing. *Magn Reson Insights* 2016; 8(Suppl 1): 69-80.

Digital Medicine and Wearable Technology. An essential part of digital medicine and wearable devices is the data mining of the incoming data for anomaly detection, prediction, and diagnosis/ decision making ([87]). The data mining process for wearable data (see Figure) includes a feature extraction/selection process for modeling/learning to yield detection, prediction, and decision making for the clinician. Expert



knowledge and metadata can influence modeling and learning.

The advent of wearable devices and sensors to continuously track physiologic parameters can provide an overall patient care strategy that will improve outcome and lower healthcare costs in cardiac patients with heart failure ([88]). This new paradigm of cardiovascular disease management can also improve the physician-patient relationship. Machine learning algorithms have also been applied to large-scale wearable sensor data in neurological disorders such as Parkinson's disease to significantly improve both clinical diagnosis and management ([89]). This sensor-based, quantitative, objective, and easy-to-use system for assessing Parkinson's disease has potential to replace traditional qualitative and subjective ratings by human interpretation.

[87] Banaee H, Ahmed MU, and Loutfi A. Data Mining for Wearable Sensors in Health Monitoring Systems: A Review of Recent Trends and Challenges. *Sensors (Basel)* 2013; 13(12): 17472-500.

[88] Steinhubl SR and Topol EJ. Moving from Digitalization to Digitization in Cardiovascular Care: Why Is It Important, and Why Could It Mean for Patients and Providers? *J Am Coll Cardiol* 2015; 66(13): 1489-1496.

[89] Kubota KJ, Chen JA, and Little MA. Machine Learning for Large-Scale Wearable Sensor Data in Parkinson's Disease: Concepts, Promises, Pitfalls, and Features. *Mov Disord* 2016; 31(9): 1314-26.

Robot Technology and Virtual Assistants. Surgical robotics such as the da Vinci system has penetrated even community hospitals and has advanced to include 3D visualization. Meanwhile, other uses of robotic technology in healthcare include delivery, sterilization, and physical therapy in various venues. Human-robot interaction is being evaluated and utilized in a variety of clinical scenarios such as rehabilitation and education ([90]). There is an ongoing debate about robotics and its ethical implications in the future of society including an exacerbation of healthcare disparities and creation of new ones ([91]).

In children with autism spectrum disorder, a robot-based approach showed that children with this disorder are more engaged in the several learning task and seem to enjoy more the task when interacting with the robot compared with the interaction with the adult ([92]). Electromechanical and robot-assisted arm training for improving activities after a stroke showed improvement of these activities in randomized controlled trials ([93]).

Much is expected of robots and virtual assistants in the future for both physical rehabilitation and psychiatric therapy as well as healthcare education and chronic disease management. If accompanied by robust AI tools, these supportive services will be particularly useful in delivering value for the patients.

[90] Sheridan TB. Human-Robot Interaction: Status and Challenges. *Hum Factors* 2016; 58(4): 525-532.

[91] Russell S, Hauert S, Altman R et al. Robotics: Ethics of Artificial Intelligence. *Nature* 2015; 521(7553): 415-418.

[92] Costescu CA, Vanderborought B, and David DO. Reversal Learning Task in Children with Autism Spectrum Disorder: A Robot-Based Approach. *J Autism Dev Disord* 2015; 45(11): 3715-25.

[93] Mehrholz J, Pohl M, Platz T et al. Electromechanical and Robot-Assisted Arm Training for Improving Activities of Daily Living, Arm Function, and Arm Muscle Strength After Stroke. *Cochrane Database Syst Rev* 2015; (11): CD006876.

Artificial Intelligence in Medicine with Focus on Selected Subspecialties

The Perfect Storm. The physicians are facing the perfect storm: exponentially increasing medical knowledge, more patients with higher degree of complexity of chronic diseases with increasingly more data, and high level of stress and burnout from the mounting burdens of EHR and workload.

There is a myriad of reasons that physicians in any subspecialty could benefit from incorporation of AI into their practices. First, the amount of medical knowledge is exponentially increasing and doubling at a rate of a few months, and yet physicians do not have enough time to read and maintain their knowledge capacity. AI can be a useful knowledge "partner". Second, AI can help organize and facilitate the care of chronic diseases in many of the patients especially as they have more relevant data from disparate sources such as genomic sequencing and wearable technology. Lastly, physicians have currently a high rate of stress and many are facing or having had burnout from their careers. The use of AI can mitigate the EHR burden and simplify their workload.

In a PubMed search for reports for artificial intelligence and medicine yielded 11,750 (382/1,126) and almost equal number of 12,790 in surgery (527/521) since 1950. Of note, there are some reports that are focused on robotics under surgery and surgical subspecialties.
AI-related reports in various subspecialties since 1950 (see Table) are separated into "high', "medium", and "low" use groups. The subspecialties range from almost 10,000 reports in Oncology (not surprising given its focus on genomic sequencing data and drug discovery), to about 5,000 in medical image-intensive fields (such as Pathology and Radiology), and more than 1,000 articles in data-intensive fields (such as Epidemiology and Critical Care Medicine). Interestingly, Neurology and Neurosurgery are on this list of top subspecialties that have an AI presence as there is much emphasis of neuroscience currently in artificial intelligence. Finally, most subspecialties, even in low and medium use subspecialties such as Cardiology, have seen an increase of at least 50% in number of published reports from 2006 to 2016. Of note, there is an incredulously low of 27 AI-related publications in Neonatology (astonishing especially since it is a data-focused specialty in a critical care setting).

Subspecialty	Total Published Reports (since 1950)	Published Reports 2006	Published Reports 2016
High Use (>1,000)			
Oncology	9,998	411	660
Pathology	6,354	224	459
Radiology	4,541	255	264
Epidemiology	2,595	99	194
Neurosurgery	1,011	74	41
Critical Care Medicine	1,006	41	66
Medium Use (250-1,000)			
Psychiatry	859	25	120
Nuclear Medicine	809	43	77
Internal Medicine	745	28	67
Neurology	715	20	110
Obstetrics and Gynecology	669	8	51
Population Health	650	27	80
Cardiology	447	18	52
Pediatrics	364	8	58
Anesthesiology	351	6	27
Genomic Medicine	301	8	78
Family Medicine	289	11	39
Infectious Disease	265	6	39
Low Use (<250)			
Ophthalmology	214	8	24
Emergency Medicine	203	4	28
Dermatology	140	4	21
Orthopedic Surgery	133	6	14
Endocrinology	120	2	16
Neonatal Medicine	107	3	16

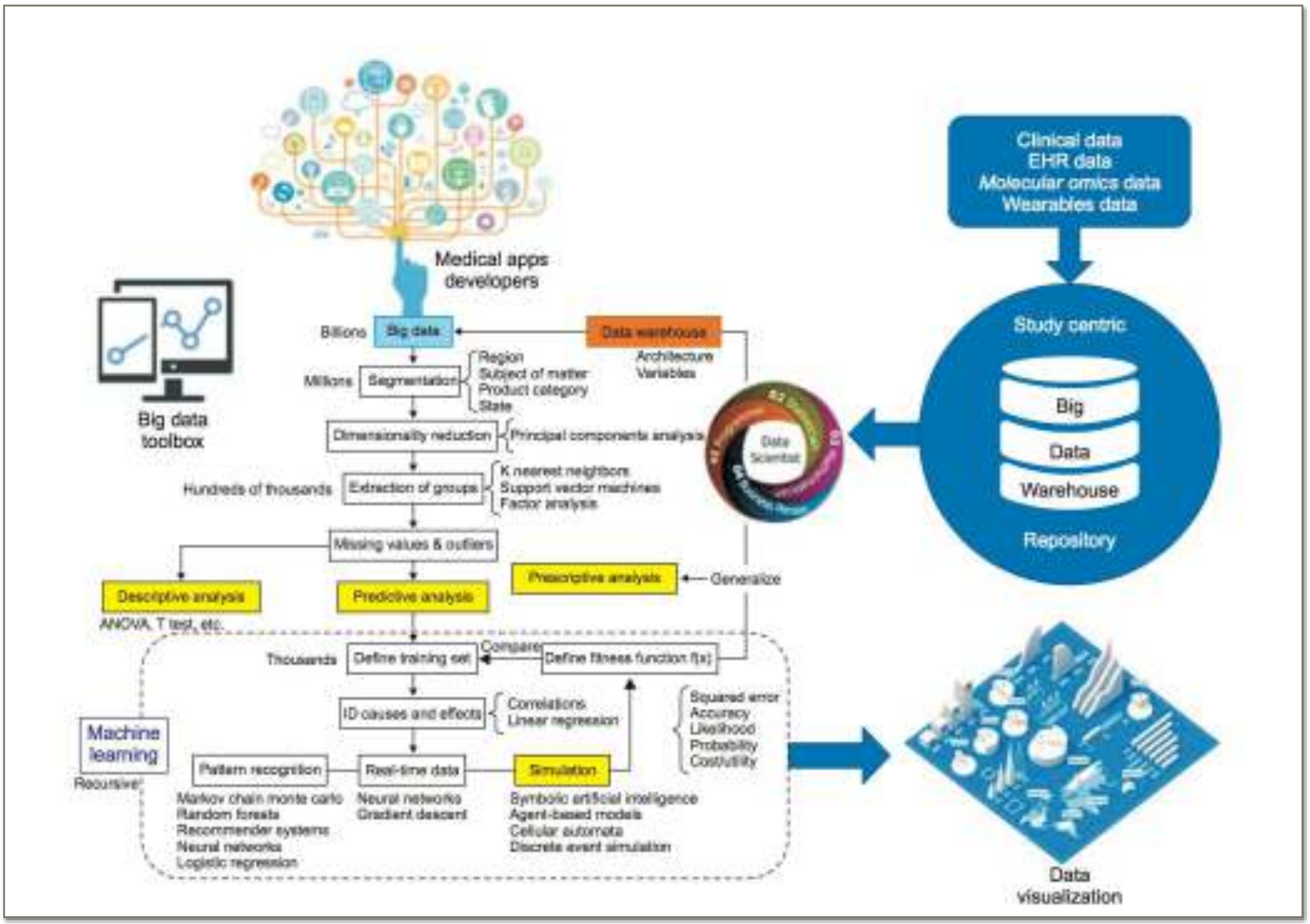
Digital Medicine and Health. Digital medicine and health herald the era of technological advances such as apps, wearable technology and remote monitoring, telemedicine and communication tools, and other diagnostic devices to affect a more optimal quality of care as well as a more timely response to any situation.

AI Strategy. The overarching theme in digital health and medicine in the use of AI is orchestrating, storing, and interpreting the huge amounts of data derived from the devices to facilitate acute and chronic disease diagnosis and management via AI-enabled acquisition and interpretation of data. This strategy will both increase the ability to proactively intervene when appropriate as well as decrease the burden on both the patient and the caretakers when the decisions are relatively straightforward.

Published Works. There is a paucity of reports in digital medicine and AI that clearly demonstrates not only proof of concept in applying AI to an app or device but also clinical benefit. As a matter of fact, a recent editorial in *Lancet* cautions the use of AI in digital medicine and strongly recommends a continual evaluation of digital health interventions for both clinical effectiveness and economic impact ([94]).

[94] Is Digital Medicine Different? Editorial in *The Lancet* 2018; 392:95.

A more positive review by Fogel discussed how AI in digital medicine can improve not only basic health screening and prevention as well as medication adherence but also the human-to-human experience of healthcare ([95]). Another review in this domain focused on the concept of a medical internet of things (mIoT) in digital healthcare that is imbued with AI-related tools ([96])(see Figure). In order to reduce overall costs for both prevention and management of chronic diseases, devices are needed to execute this strategy: to monitor health biometrics, to auto-administer therapies, and to track real-time health data during therapy. Along with these devices, mobile applications for access to medical records as well as tools for telemedicine and telehealth for this new paradigm of medical IoT. All of these devices and equipment will need an AI-centric strategy for data integration and interpretation for delivering optimal

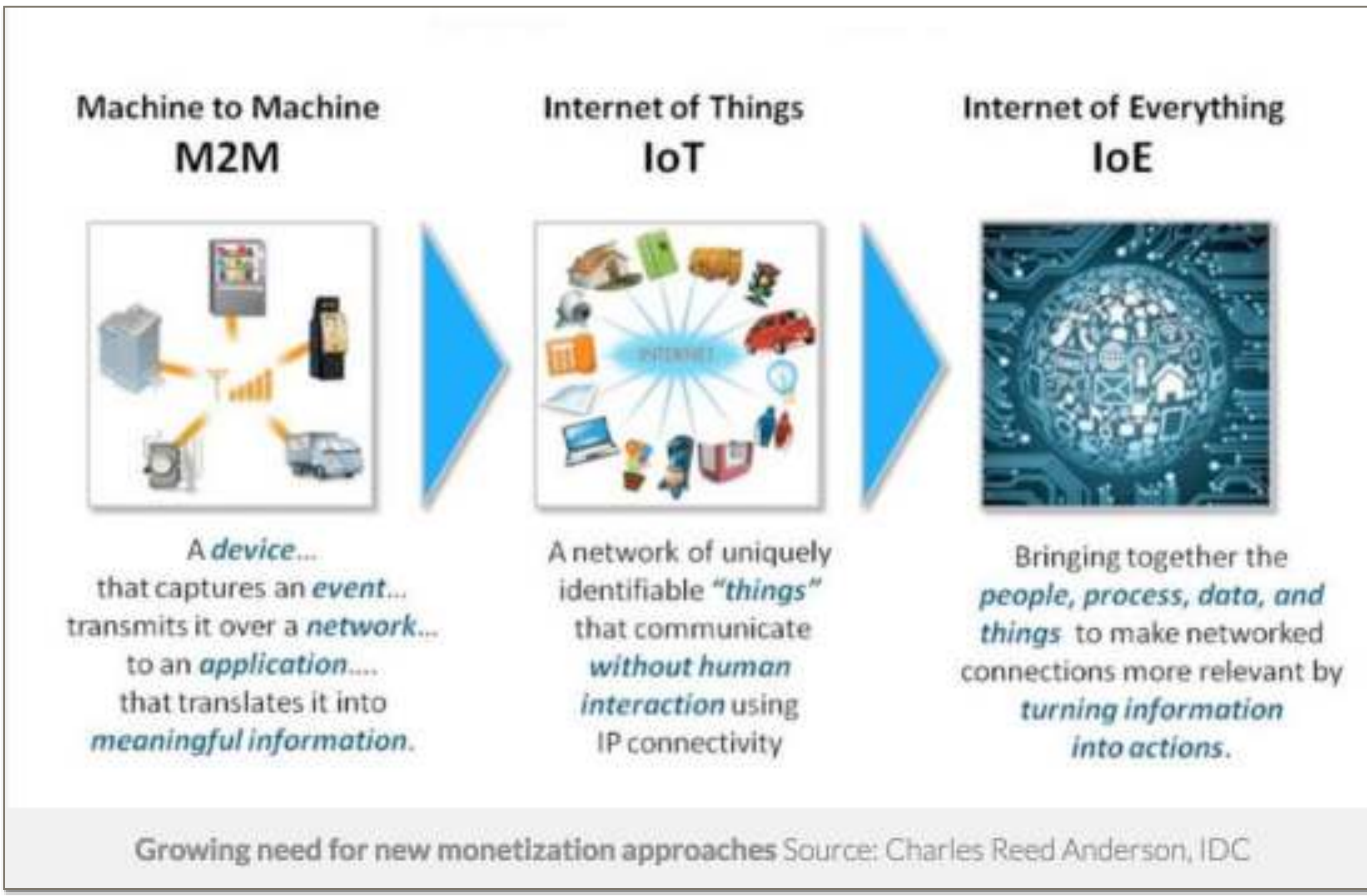


healthcare advice and direction.

[95] Fogel AL and Kvedar JC. Perspective: Artificial Intelligence Powers Digital Medicine. *npj|Digital Medicine* 2018; 1:5-8.
[96] Dimitrov D. Medical Internet of Things and Big Data in Healthcare. *Healthc Inform Res* 2016; 22(3): 156-163.

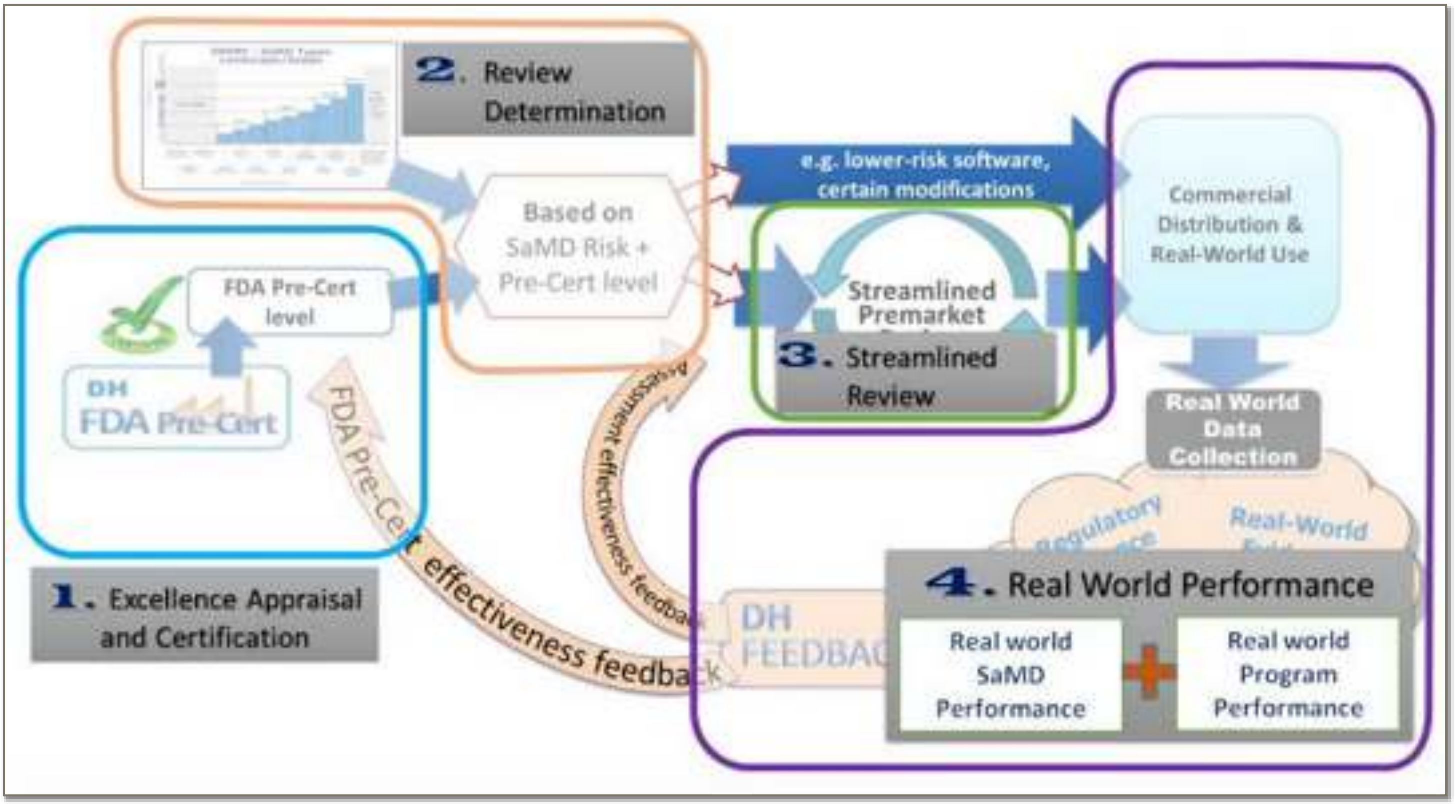
While chronic diseases such as diabetes care can benefit greatly from a coordinated and efficient strategy, use of technology including AI remains fragmented at present due to a myriad of issues: lack of supportive policy and regulation, unsustainable reimbursement, inefficient business models, and concerns regarding data security and privacy ([97]).

Future Applications. In the near future, embedded AI (eAI) and machine learning algorithms evolve toward the internet of everything (IoE) and will bring together people, process, data, and things (see Figure); this strategy will allow the accrued data be streamlined and organized in the cloud proactively in an overall paradigm of personalized precision medicine. As these devices become more intelligent, increasingly higher levels of sophistication in decision support can also be part of both 1) preventive medicine (such as retinal images for retinopathy screening or skin lesions for melanoma detection) as well as 2) chronic disease care management (such as diabetes, hypertension, or heart failure).



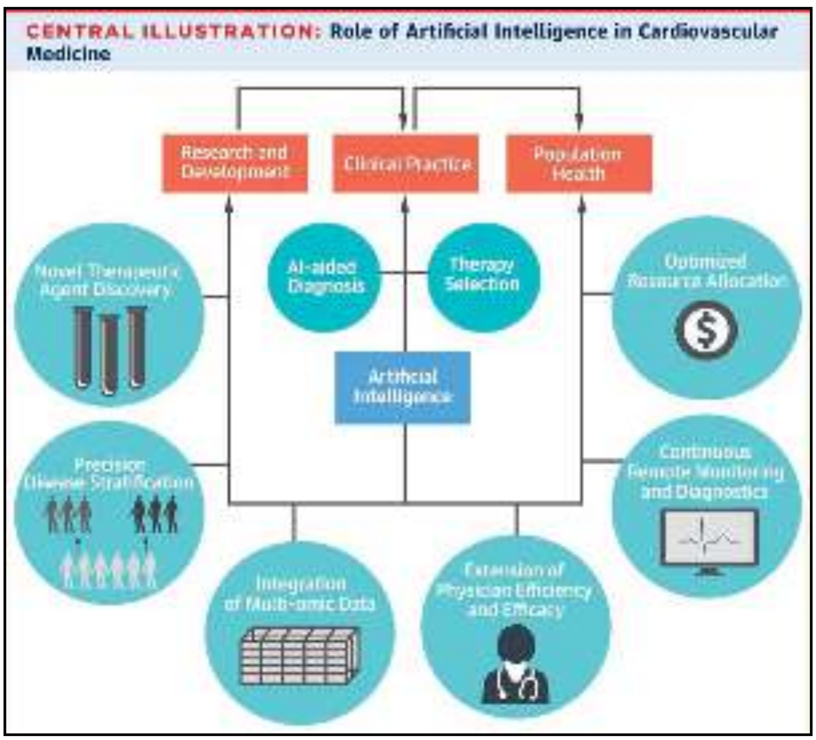
[97] Fatehi F, Menon A, and Bird D. Diabetes Care in the Digital Era: A Synoptic Overview. *Curr Diab Rep* 2018; 18(7): 38-47.

An overall strategy for preliminary and continual evaluation of AI applications in digital medicine is needed as the barrier to entry may continue to be low for some apps and devices. This evaluation process will need insight from not only organizations such as the AMA or the FDA (see Figure), but perhaps also by an international consortium of multidisciplinary experts. Finally, attention needs to be directed towards the cybersecurity of these intelligent devices to mitigate the risk of data breaches and therefore intentional harm to patients and caretakers.



Cardiology (Adult and Pediatric). Cardiology is a subspecialty that deals with disorders of the heart and blood vessels: while an adult cardiologist sees adults with these disorders (coronary artery disease, congestive heart failure, elevated cholesterol, atrial fibrillation, or hypertension), a pediatric cardiologist treats children and adolescents with both acquired and congenital heart disease. Typical tests that a cardiologist orders include: electrocardiogram (EKG), echocardiogram, CT or MRI of the heart and chest, nuclear imaging, exercise testing, and tests that deal with heart rhythm disturbances much as a Holter monitor. Cardiologists also perform cardiac catheterizations for diagnostic purposes as well as for interventional procedures such a balloon angioplasty, stent implantation, or electrophysiologic study or a pacemaker insertion. A cardiologist may see patients both in the hospital setting (including often in the ICU) or in the outpatient clinic.

AI Strategy. As cardiology is both a perceptual or image intensive field as well as a cognitive or decision making subspecialty, AI is a particularly useful technology for cardiology. In the medical image domain, deep learning with CNN can be deployed for not only static images such as EKG, CT and MRI but also dynamic images such as angiograms and echocardiograms. In the complex decision making area, the use of reinforcement learning can be particularly useful for the ever increasingly complex nature of diagnostic and therapeutic precision cardiovascular medicine.



Published Works. A recent review by Johnson et al in *Journal of the American College of Cardiology (JACC)* delineates how AI will affect all aspects of cardiology: from research to clinical practice and population health (see Figure)([98]). An earlier useful review focused on how effective AI can be helpful in the context of precision cardiovascular medicine ([99]).

[98] Johnson KW et al. Artificial Intelligence in Cardiology. *Journal of the American College of Cardiology* 2018; 71(23): 2668-2679.
[99] Krittanawong C et al. Artificial Intelligence in Precision Cardiovascular Medicine. *Journal of American College of Cardiology* 2017; 69(21): 2657-2664.

There are increasing use of AI and data science in the domain of cardiac imaging for both diagnosis and prognosis. Deep learning algorithms have been applied to cardiac MRI as a prognosis prediction tool in patients with pulmonary hypertension and shown to be superior to clinicians assessment ([100]). In addition, machine learning was applied to patients with preserved ejection fraction heart failure and helped to set up a new phenotypic system for heart failure ([101]). Finally, arrhythmia detection with a single-lead ECG at a cardiologist level has been accomplished with a 34-layer CNN ([102]).
There are also reports of using AI for clinical decision support in various settings. Decision making in cardiology is often complex and vulnerable to many heuristics and biases ([103]). For estimating risk in congenital heart surgery, four AI-based algorithms were employed to facilitate a clinical decision support system ([104]). One report described using machine learning and system modeling to facilitate a multi center collaborative learning project for rapid structured fact-finding and dissemination of expertise (). This forward thinking approach can provide a complement to the traditional randomized clinical trials ([105]).

Future Applications. In the near future, areas for AI application in cardiovascular medicine are numerous and can include 1) searching for new phenotypic and "image"omic expressions of common cardiovascular diseases such as heart failure, 2) applying augmented and virtual or even mixed reality applications for education, training, and pre-procedure trials, and 3) using reinforcement learning combined with cognitive architecture for complex precision cardiovascular medicine diagnosis and management.

---

[100] Dawes TJW, de Marvao A, Shi W et al. Machine Learning of Three-Dimensional Right Ventricular Motion Enables Outcome Prediction in Pulmonary Hypertension: A Cardiac MR Imaging Study. *Radiology* 2017; 283(2): 381-390.
[101] Shah SJ, Katz DH, Selvaraj S et al. Phenomapping for Novel Classification of Heart Failure with Preserved Ejection Fraction. *Circulation* 2015; 131(3): 269-279.
[102] Rajpurkar P, Hannun AY, haghpanahi M et al. Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks. arXiv:1707.01836v1[cs.CV] 6 July 2017.
[103] Ryan A, Duignan S, Kenny D et al. Decision Making in Pediatric Cardiology: Are We Prone to Heuristics, Biases, and Traps? *Pediatric Cardiology* 2018; 39(1): 160-167.
[104] Ruiz-Fernandez D, Torra AM, Soriano-Paya A et al. Aid Decision Algorithms to Estimate the Risk in Congenital Heart Surgery. *Computer Methods and Programs in Biomedicine* 2016; 126:118-127.
[105] Wolf MJ, Lee EK, Nicolson SC et al. Rationale and Methodology of a Collaborative Learning Project in Congenital Cardiac Care. *American Heart Journal* 2016; 174:129-137.

Problems and Potential Solutions

Extreme uncertainty and dissatisfaction in data and information exists in the imbroglio of the practice of medicine and the world of healthcare especially with the increased incorporation of electronic health records (EHR) into the hospital and clinics and limited input into the few EHR vendors. Some of the data issues facing the present day practitioner include: escalating and missing data, exponentially increasing information, rigid regulatory policies, and decreasing access to information.

In addition, medical providers lack sufficient insight and education in the realm of data science and this ignorance results in an inadequate knowledge from the rich data that now exist in healthcare and medicine. To date, there has not been any publications on deep learning in *Lancet* or *The New England Journal of Medicine* nor has there been wide acceptance of data science or artificial intelligence in the medical school or clinical training program curriculum.

Finally, there is also a significant cultural and intellectual schism between the clinical world and data science domain: most medical meetings lack data science or artificial intelligence discussions on application and gatherings of devotees of machine learning or data science in healthcare or medicine also lack strong clinician presence.

Practitioners require accurate data as well as up to date information and sharing of ideas to ensure best outcomes for the patient population([106]). In the future, the data domain in healthcare will also be more inclusive of non-traditional sources such as social media and home monitoring especially with the proliferation of applications.

[106] Darcy AM, Louie AK, and Roberts LW. Machine Learning and the Profession of Medicine. *JAMA* 2016; 315(6): 551-552.

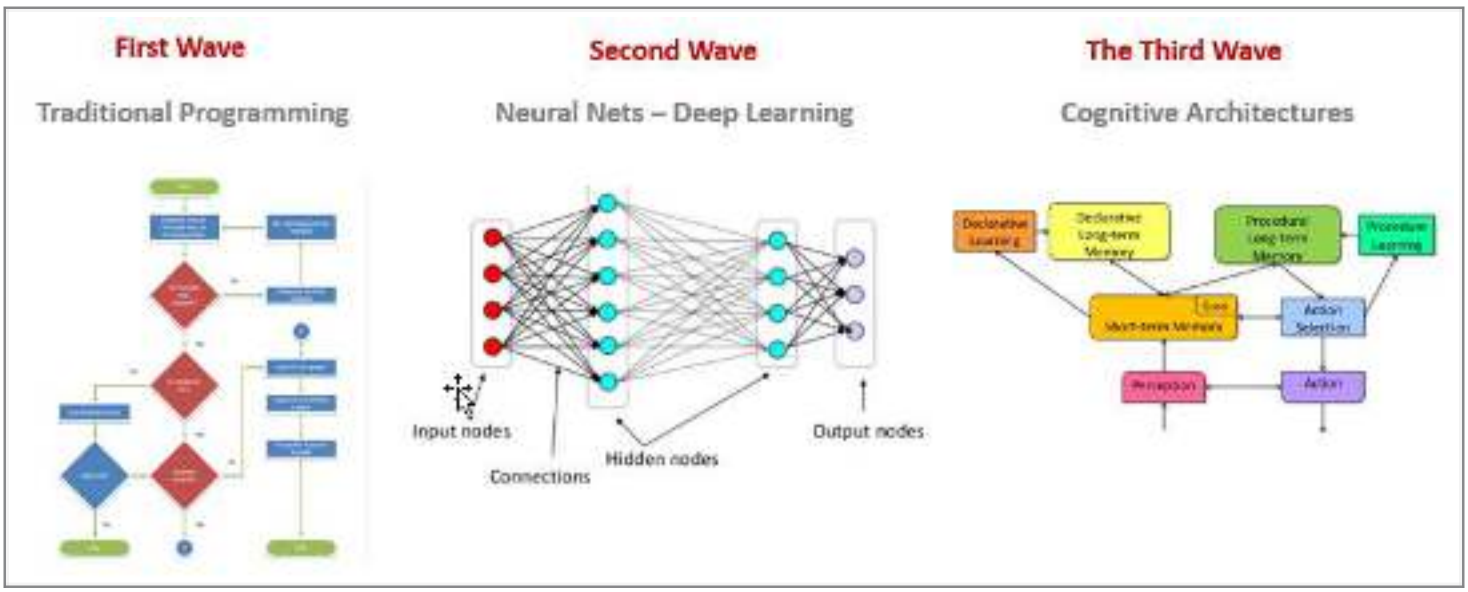
In conclusion, the convergence of Big Data, improved algorithms, computational power, and cloud storage in healthcare has started to yield robust machine learning projects and reasonable results in biomedicine and healthcare. As deep learning usually mandates large training sets and the medical field often lack such large-sized databases, it behooves the medical field to be better organized and more collaborative in such endeavors and concomitantly for the data science stakeholders to accommodate this limitation with innovative approaches in deep learning. Best practice in the near future will involve the use of data and analytics to answer the clinical questions (intelligence-based medicine) rather than current practice of solely relying on published reports (evidence-based medicine).

IV. THE FUTURE OF ARTIFICIAL INTELLIGENCE & APPLICATION IN MEDICINE

Introduction

The future of artificial intelligence is very promising. This premise is based on rapid improvement and accelerated deployment of advanced computer technology (such as quantum computing, neuromorphic chips, and cloud storage) as well as rapid development and evolution of artificial intelligence techniques (such as deep learning and it many variants) that has lead to a Cambrian explosion of AI.

Peter Voss so aptly described the future of AI as the “third wave”: the first wave of AI was good old fashioned AI (GOFAI) that focused on traditional programming followed by the second wave of AI of the current deep learning so that this third wave will be reliant on many cognitive architectures (see Figure). While this third wave of cognitive architecture is much more complex as it incorporates the relevant elements of human cognition, the former two approaches do lack the biological completeness and integration of the third wave. Finally, in this third wave, cognitive architectures will need to possess: general learning ability; real-time and interactive learning; dynamic goals and context; transfer learning;

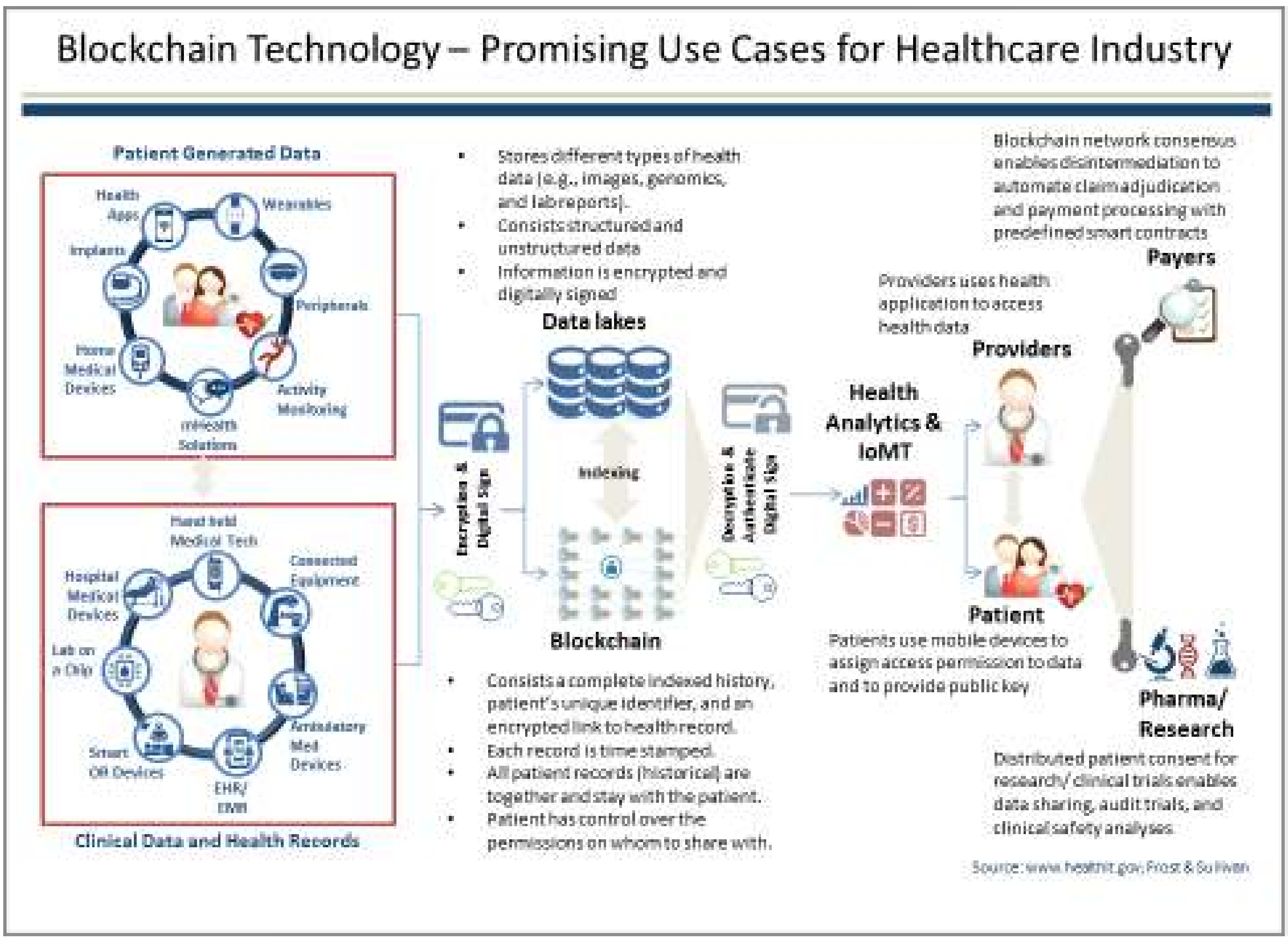


and abstract reasoning in order to reach AGI.

The Future of Artificial Intelligence

Augmented and Virtual Reality. The future of AI will enable both augmented and virtual reality. Augmented reality (AR) is an enhanced reality that is a result of computer-generated enhancements atop of reality. Virtual reality (VR), coded by a special language called virtual reality modeling language (VRML), is a computer-generated artificial simulation or recreation of a situation mainly via vision and hearing (Facebook Oculus is an example). Both of these AI-enabled reality methodologies will become commonplace in the near future.

Blockchain and Cybersecurity. Block chain (or blockchain) is the use of cryptography to allow a collection of blocks or records to be maintained in such a way that is difficult to modify. This strategy was initially used for bitcoin as a public ledger. Blockchain is therefore a disruptive innovation in information registration that utilizes three existing technologies: private key cryptography, peer-to-peer (P2P) network, and the blockchain protocol. A successful deployment of technologies such as blockchain to



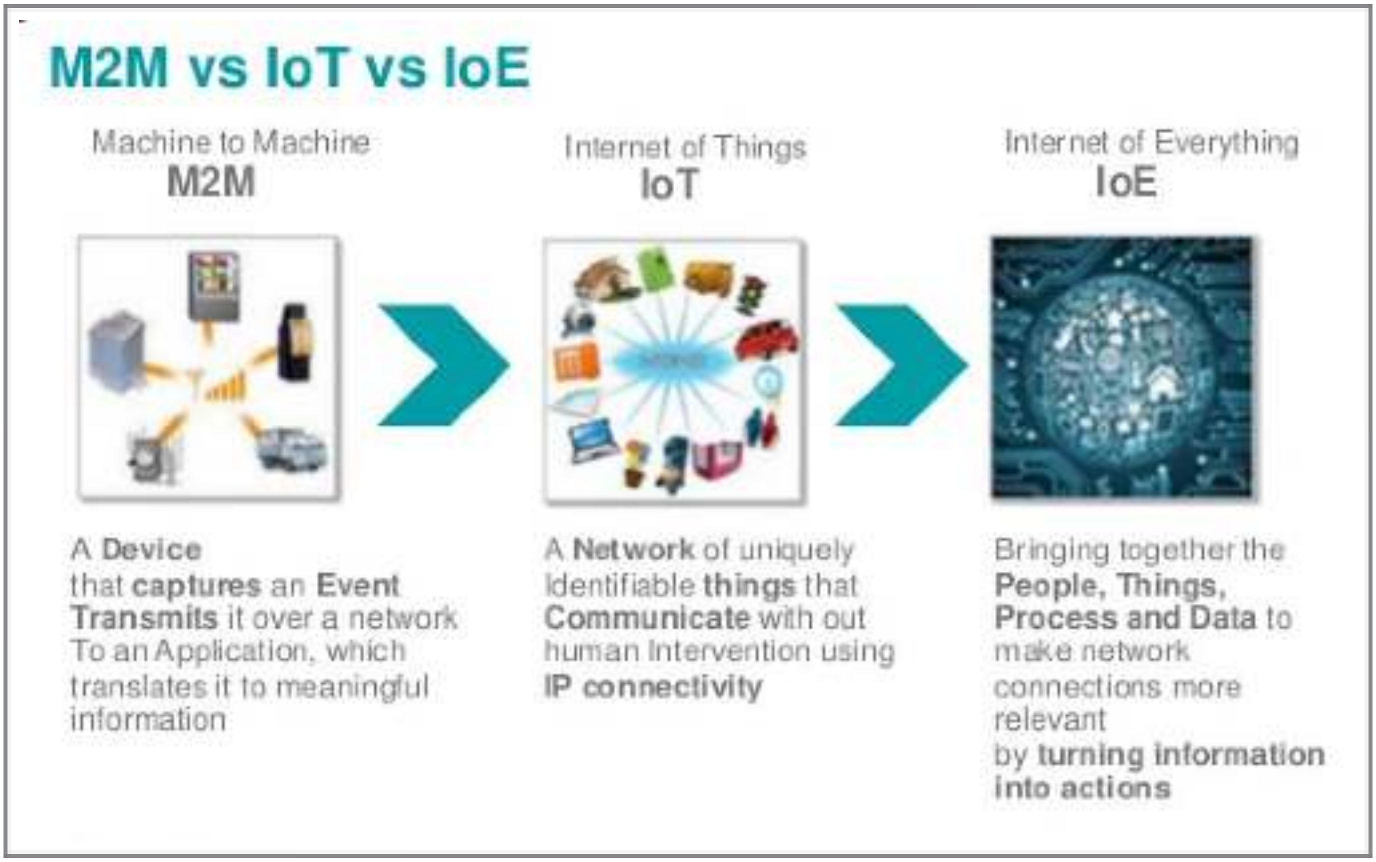
improve cybersecurity in healthcare will facilitate data sharing amongst stakeholders in the near future.

Cloud Computing. Cloud computing has been thus far a public cloud model (exemplified by Amazon Web Services (AWS) or Salesforce's CRM system), but the cloud of the future will enable virtualization and management of software-defined data services. Future AI applications, therefore, will be in the cloud so AI will be in the form of AI-as-a-service (AIaas). The other aspect of cloud computing that will be essential for AI is its role in the formation of internet of everything (IoE)(see below) to create a ubiquitous decentralization of devices and sensors.

Deep Learning (and Other Types of Learning). Deep learning seems to be evolving more rapidly as time goes on. Reinforcement learning has its origin over 100 years ago with the psychologist Edward Thorndike and his cat experiments in which the cats learned to "reinforce" their positive outcome with the appropriate behavior of pushing a lever. In 1951, Harvard's Marvin Minsky devised an equipment to emulate this nature's reinforcement learning; this equipment, called Stochastic Neural Analogy Reinforcement Computer (SNARC), consisted of motors, tubes, and clutches that functioned as dozens of neurons and synapses that favored behavior that lead to a positive outcome. Reinforcement learning and its wondrous capability was best demonstrated in the Google DeepMind's *AlphaGo* program and its recent successful defeat of the Go champion Lee Se-dol. The real dividend in reinforcement learning is in its potent combination with deep learning, which uses a large neural network for pattern recognition, in the form of deep reinforcement learning.

There is also transfer learning that occurs when a network that is trained for one task is then used to configure the network for another task. In addition, according to Yann LeCun, director of Facebook AI, generative adversarial networks (GANs) is "the most interesting idea in the last ten years in machine learning." In GANs, two adversarial models (called "generator" and "discriminator") can co-train through back propagation as a form of unsupervised learning (thus giving the computer the capability of "imagination"). Finally, Geoff Hinton of Google has very recently introduced the concept of "capsule" neural nets, a neural network that found its inspiration in the neurons in our brains and therefore perhaps be "smarter" than regular neural networks in that it will need less input data to perform. Overall, these recent developments seem to indicate a convergence of deep learning and cognitive approaches to AI (perhaps "deep cognition"?).

Internet of Everything (IoE). Data in health care will concomitantly escalate as well from the advent of wearable and home monitoring technology leading to Internet of Everything (IoE)([107]). There are now an estimated 100,000 mHealth apps available ([108]). While the Internet of Things (IoT) is the interconnection of billions of physiologic devices to the Internet, IoE will be essentially a "network of networks" to incorporate people, processes, and data to these devices to enable automation in data acquisition and analytics without human intervention ([109]). All of these "smart" devices with wireless sensor networks (WSNs) will add to the collective intelligence of medical data and information. The internet of everything is basically intelligent connection of people, process, data, and things (see Figure); it builds on IoT by adding network intelligence and turning information into actions. The future of AI will need to involve the IoE just as the brain needs a nervous system.



The E could also be for embedded AI, or eAI.

[107] Banaee H et al. Data Mining for Wearable Sensors in Health Monitoring Systems: A Review of Recent Trends and Challenges. *Sensors* 2013; 13:17472-17500.

[108] Cortez NG et al. FDA Regulation of Mobile Health Technologies. *N Eng J Med* 2014; 171(4): 372-379.

[109] Gubbi J et al. Internet of Things (IoT): A Vision, Architectural Elements, and Future Directions. Future Generation Computer Systems 2013; 29: 1645-1660.

Other Key Concepts

Virtual Assistants and Bots. A virtual assistant is an AI-inspired software agent that is capable of performing certain tasks or services via text or voice. Examples are Apple's Siri, Google Assistant, and Amazon's Alexa. A chatbot (or bot) is a service that is capable of conducting a conversation with a human as a result of using rules governed by AI. Sophisticated chatbots can even use machine learning and therefore can get "smarter" as it converses with people. Other names for this entity include: virtual or conversational agent.

Swarm Intelligence. This is intelligence derived from many individuals based on self-organizing group behavior. The collective behavior illustrates that unified systems outperform the majority of individual members, but since humans do not naturally have these connections as observed in ants or fish, swarm AI is executed by technology to provide feedback to human members. In short, this type of intelligence leverages the "wisdom of the crowd".

Quantum Computing. This futuristic computing is a new approach to process information and is much more powerful than the conventional computer. A quantum computer utilizes quantum bits (or qubits) instead of the conventional bits (that is in 0 or 1 states) used in digital computing. Quantum computing takes advantage of the quantum phenomenon that subatomic particles can exist in more than one state at a time. A quantum computer like the D-Wave computer is 100 million times faster than a conventional laptop.

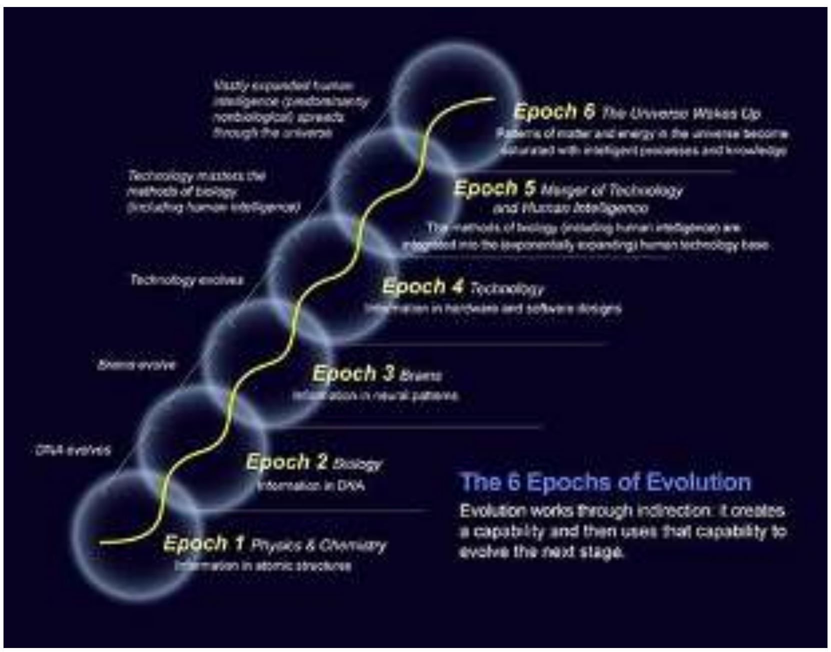
Neuromorphic Computing. Also known as neuromorphic engineering, this is a concept in which computer chips can mimic the brain by communicating in parallel using "spikes", which are bursts of electric current that the neuron can control. These neuromorphic chips have the advantage over traditional computer CPUs in that these chips require far less power to process AI-inspired algorithms.

Hypergraph Database. Traditional relational databases are weak in complex hierarchical data and processing graphical data structures. A graph database is designed to neutralize these disadvantages by processing data in a graphical strategy and enable queries across the data network. In order to model even more complex and highly interconnected data, a new paradigm of data representation called hypergraphs will need to be implemented. A hypergraph is a graph model in which the relationships (called a hyperedges) can connect any number of nodes.

Brain-Computer Interface (BCI). The other terms include mind-machine interface (MMI), brain-machine interface (BMI), or direct neural interface (DNI). These are all communication pathways between the brain and an external device to augment natural intelligence. An example of this type of device is the one proposed by Elon Musk called "neural lace".

A myriad of issues in use of artificial intelligence remain. One is the ethics of its use in the variety of sectors and the accompanying debates amongst scholars and scientists as well as the public. Elon Musk and Stephen Hawking both predict dire consequences while other Silicon Valley titans argue the other way. The truth may very well be in the middle: we need to be respectful of the power of AI and not be careless in its deployment. One approach is that suggested by Oren Etzioni to have three rules for AI that are inspired by Isaac Asimov's three rules for robotics: 1) An AI system must be subject to full gamut of laws that apply to its human operator; 2) An AI system must clearly disclose that it is not human; and 3) An AI system cannot retain or disclose confidential information without explicit approval from the source of that information ([110]).

Another issues is the economics of such a paradigm shift in work and compensation. Artificial intelligence is a prediction technology so cost of goods and services that rely on prediction (such as inventory management and demand forecasting) will fall. But since all human activities rely on not just data and prediction but also judgment, action, and outcome, prices for the latter three can increase as the demand for these three capabilities go up.



The issue of singularity (in which the computers will be more intelligent than all of mankind and replace its intellectual capacity) and how we accommodate this epoch (see Figure) will need to be discussed. The economic impact of such an event especially with countries which do not have large investments in AI needs to be examined.

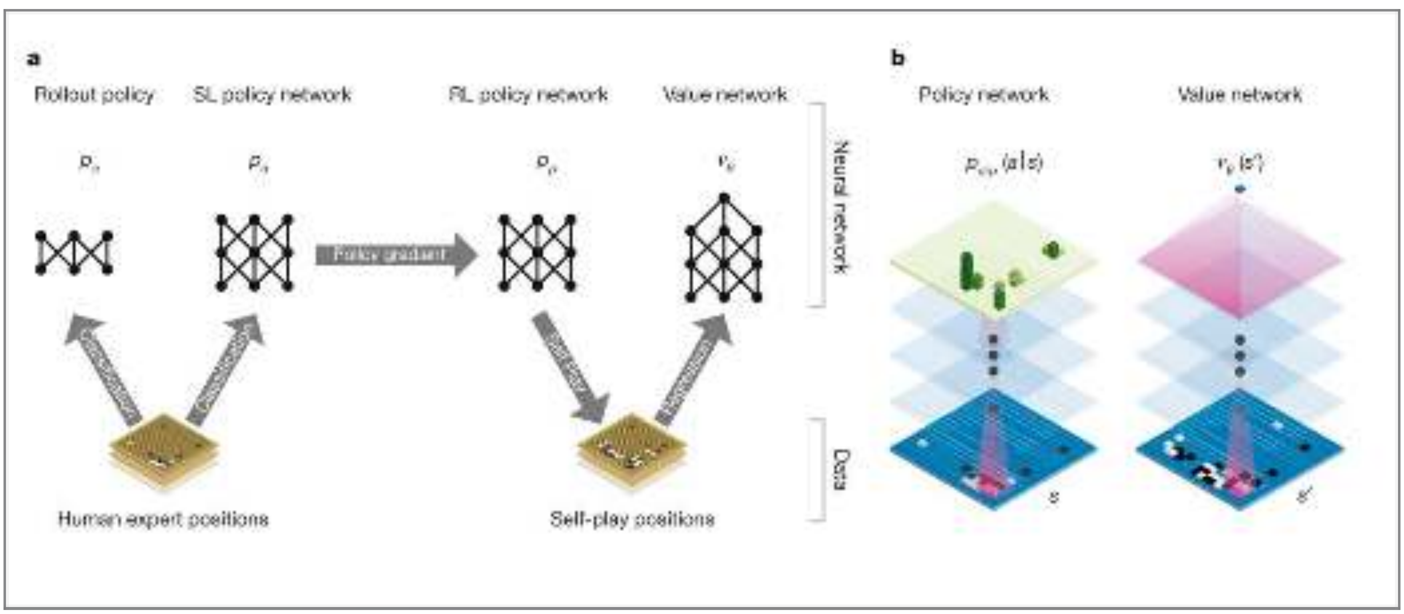
Lastly, the issue of bias in the formation and application of AI for all uses will need to be scrutinized. Biased date feeding into machine learning algorithms can lead to biased systems in AI ([111]).

[110] Etzioni, O. How to Regulate Artificial Intelligence. New York Times, September 1, 2017.
[111] Knight W. Forget Killer Robots- Bias Is the Real AI Danger. *MIT Technology Review*, 2017.

The Future of Artificial Intelligence in Medicine

The One Hundred Year Study on Artificial Intelligence (AI100) is a long-term study of AI and its impact on people and society. The eight relevant areas that are considered most salient include healthcare (the other seven being transportation, service robots, education, low-resource communities, public safety and security, employment and workplace, and entertainment)([112]).
Already used in areas such as radiology, pathology, genomic medicine, cardiology, outpatient services, and intensive care, AI will continue to have escalating impact in medicine albeit with concomitant fear amongst stakeholders for an AI "takeover", especially their jobs.

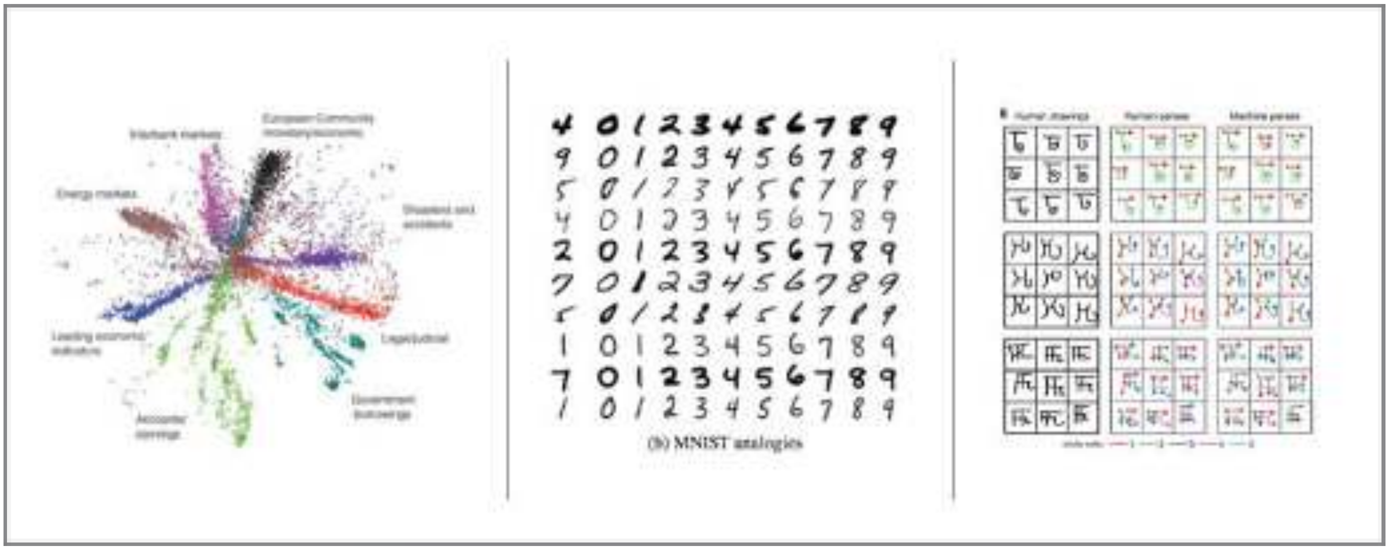
There are exciting advances in the computer science and AI realm for healthcare and medicine. Various subtypes and variations of deep learning are evolving within the machine learning paradigm: deep, reinforcement, deep reinforcement, transfer, spiking, and one-shot types of learning will figure prominently in the future of AI and especially for biomedicine as these methodologies can accommodate the nuances and complexities of medicine. *AlphaGo* is ideally designed for the myriad of nuanced humanlike decision-making aspects of medicine since it accommodates for sequences of better decisions by a combination of recognition of complex patterns, long-term planning, and "intelligent" decision-making (see Figure). *AlphaGo* has three components: 1) Policy network- this element evaluates the current situation and predicts the next step; 2) Fast rollout- this component improves the speed of the decision; and 3) Value network- this element evaluates the situation and predicts which side will win. These three components may be ideally suited for many decisions in medicine and healthcare especially



as it pertains to precision medicine and population health.

112 https://ai100.stanford.edu.

In addition, Fei-Fei of Stanford promulgated one-shot learning that can bring a special dimension to unique cases in medicine as it will not require the usual large dimensionality of data that the previous types of learning will need (see Figure where unsupervised, semi-supervised, and one-shot learning are

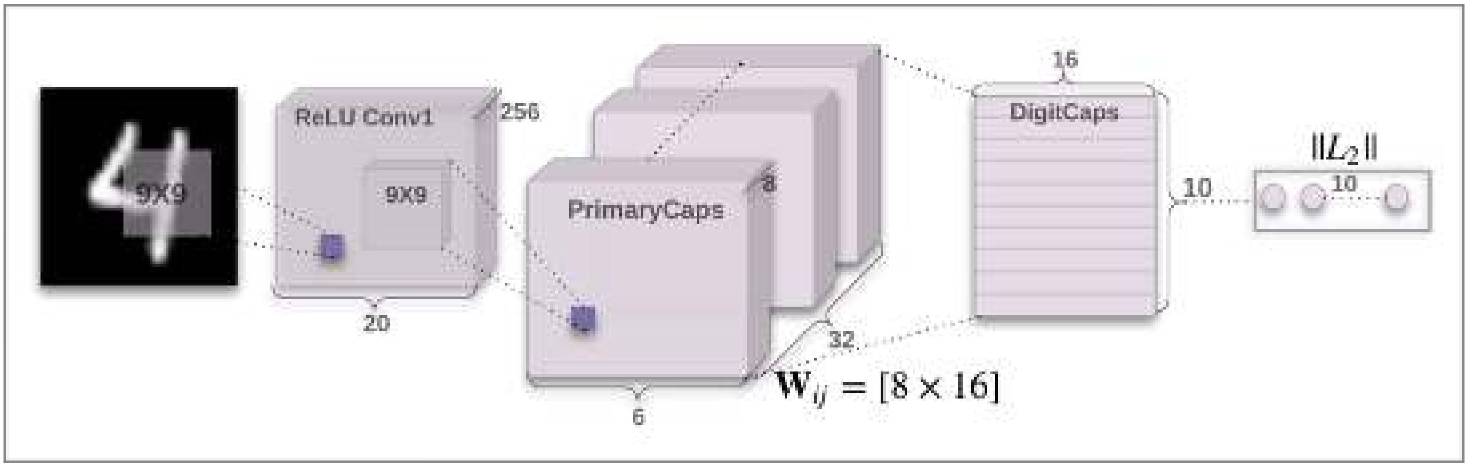


illustrated on the left, center, and right respectively)([113]).

A third-generation neural network called spiking neural network (SNN) that bridges the gap between machine learning and neuroscience uses discrete events rather than continuous values that conventional machine or deep learning use. Similarly, Hinton's recent description of a capsule network (or "capsules")(see Figure) with an element of biomimicry can be "smarter" with less input data and is potentially ideally suited for biomedical data. The conventional CNNs were instrumental in the popularity of deep learning but have significant limitations. One such drawback is the lack of spatial hierarchies

[113] Ballinger B. Three Challenges for Artificial Intelligence in Medicine. *Cardiogram*, September 19, 2016.

between the objects. Capsules essentially introduce "intuition" to deep learning as these entities



improves model hierarchical relationships inside the knowledge representation.

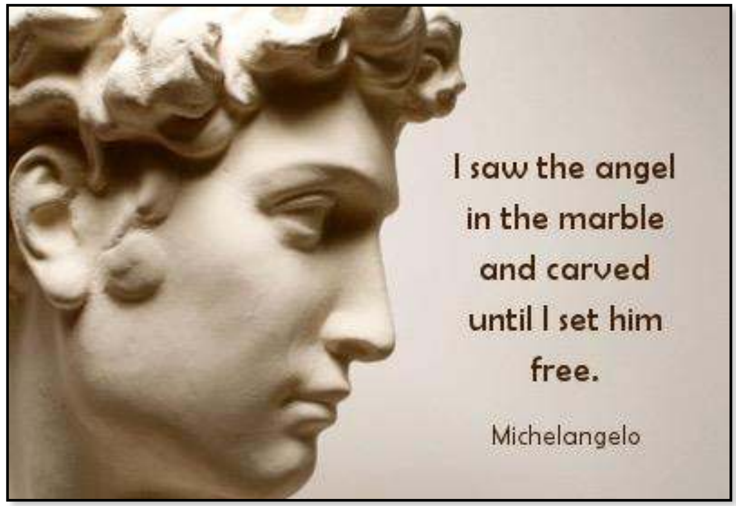
The future of biomedicine can include a proposal for an artificial intelligence-inspired cloud continuum of data-information-knowledge-intelligence (a "medical intelligence" as a service, or "MIaaS"). In Nick McKeown's parlance, medical data and analytics need to be transformed from a "vertically integrated, closed, proprietary, slow innovating" data system to that of a "horizontally integrated, open interfaces, and rapid innovation" data ecosystem ([114]). To put it into biological lexicon, the biomedical data system needs to transform from a rudimentary musculoskeletal system to an intelligent nervous system. In addition, cloud computing and storage will be vital to facilitate the panoply of AI techniques for multi-institutional collaborations that will be essential for the future of AI in biomedicine and healthcare.

By embedding intelligence into all aspects of medical data from graph database and meta-database management system to cloud infrastructure and even software-driven virtualization, the aforementioned strategies can accelerate this transformation in biomedicine from fragmented and unstructured data sets to cohesive and agile information imbued with medical intelligence.
Many of the myriad of technological advances will be essential in the evolution of AI in biomedicine. Quantum computing and neuromorphic chips are important advances that are elements that can accelerate the capabilities of AI tools especially with the vast amounts of data in medicine. Even DNA computing for storage has been suggested as a methodology to increase storage capacity. The "medical" internet of things and everything (mIoT and mIoE) will also provide the rich data sources for medicine in the form of wearable and monitoring devices from both hospital and home. Building a healthcare internet will need four essential components: 1) things (monitors); 2) gateway; 3) data storage (Hadoop)/ cloud (Amazon web services, Google cloud platform, IBM Bluemix, or Microsoft Azure); and 4) machine learning (TensorFlow, Apache Spark, R, etc). The advent of augmented and virtual reality will create limitless new opportunities for not only patient care but also training and education for medical personnel. Novel hypergraph databases as well as human swarm intelligence of many stakeholders from different institutions will improve decision-making processes. Finally, the full promise of AI in medicine will require a data security and privacy transformation (via disruptive technology such as blockchain) and an innovation focus from regulatory agencies and software vendors (via open source collaboration).

Perhaps clinicians and medical educators can embrace rather than distrust AI and allow its capacities to transform how we teach and deliver healthcare, especially in the changing healthcare milieu to fee-for-value systems. An effective AI in medicine strategy will liberate clinicians from the burgeoning burden of electronic health records and allow a return to an ideal physician-patient relationship. In addition, since computers are excellent at handling data and making predictions, the human judgment will become even more valuable.

[114] McKewon N. *Making SDNs Work*. YouTube, April, 2012.

Ten Elements for Successful Implementation of Artificial Intelligence in Medicine



A modification of the Michelangelo quote on the statue David (see Figure) can be aptly applied to the future of AI in medicine:
"I saw the intelligence in the (medical) data and mined until I set the wisdom free."
The future holds great promise for implementation of artificial intelligence in medicine. These are ten potential solutions to present problems in artificial intelligence in medicine to an ultimately successful paradigm shift:

Improved data storing and sharing strategies in healthcare. Much of the data in healthcare is missing, inaccurate, and disorganized as well as fragmented. In addition, patients need to be empowered to own their data as often healthcare data are sequestered in a hospital or clinic with virtually no access. Finally, all the stakeholders must come together and be willing to share healthcare data. If AI is the rocket to launch us into orbit and onto moonshot projects, data is the fuel we need yet this fuel is not centrally collected. Timing is key as major data sources are still yet being formed, such as genomic data, wearable technology data, and socioeconomic data. While AI methodologies are ahead of schedule, data in healthcare has been in disarray for decades and will require effort to be improved. One key strategy will be the deployment of a revolutionary data infrastructure (graph or hyper graph databases). All of this discussion of healthcare data being collected and stored will require close attention to cybersecurity and the emergence of blockchain technology may be timely. In short, good AI in medicine mandates good not merely basic data and database foundation but also some level of connectivity via IoT and IoE.

AI in medicine awareness and education. Even with escalating increases in venture capital in the area of artificial intelligence in healthcare and medicine, particularly in the areas of decision support and medical imaging, hospital administrators and clinicians as well as investors still lack sufficient education about artificial intelligence methodologies applied to healthcare and medicine. There is a paucity of data science classes in medical schools and residency programs.
In addition, computer and data scientists are also less than fully enlightened in the realm of clinical medicine and what clinicians actually need most for artificial intelligence to be utilized in order to lessen their burden. An overall higher level of knowledge will also prevent procuring AI solutions that are "Mechanical Turks".

Human-to-human (H2H) collaboration for AI-driven agenda. In the work of artificial intelligence in medicine, it is vital that human to human interactions and relationships drive these agendas and projects. From sharing data to doing projects using artificial intelligence in healthcare and medicine, human champions and leaders from a myriad of domains need to be the drivers of these joint agendas.

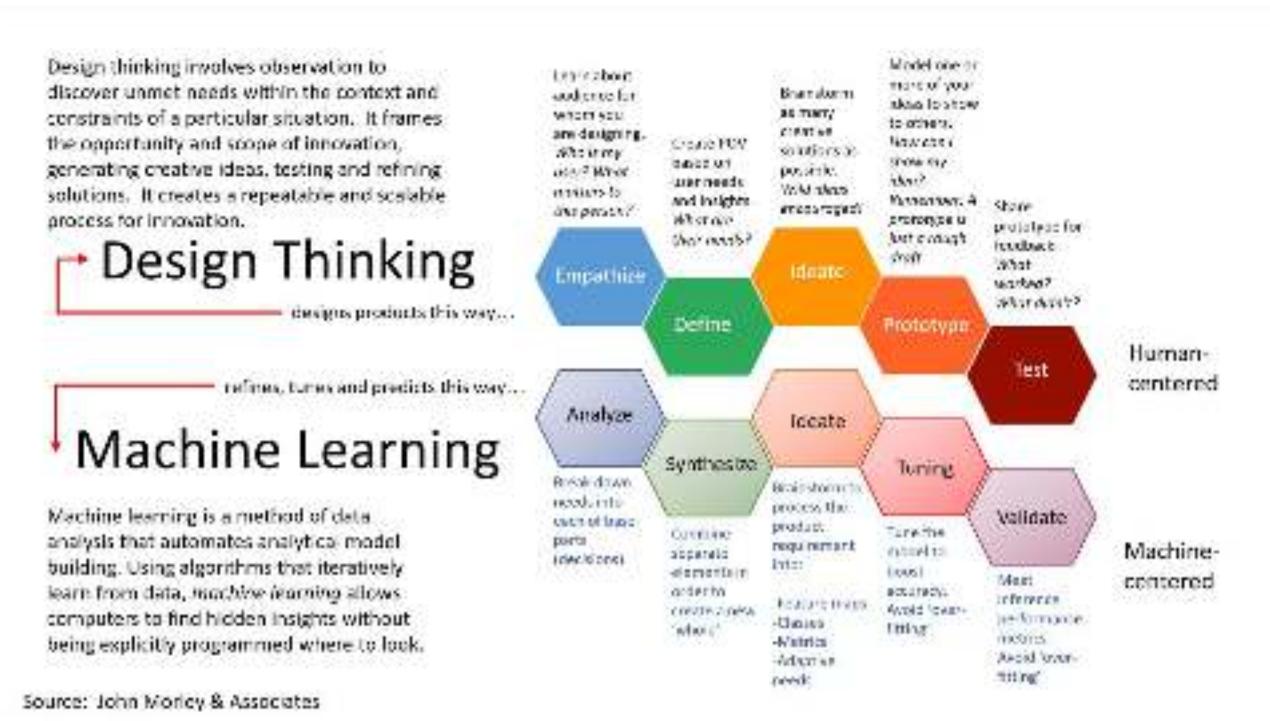
Clinician-to-data scientist synergy for AI in medicine. There is a general lack of awareness and education for clinicians about deep learning (try to find a clinically active physician who even understands what deep learning is) and concomitantly for data scientists about clinical medicine (try to find a data scientist who interfaces with a physician more than once in a while and who actually spends time in the clinical domain). This clinician-to-data scientist distance is further increased with at times human hubris on both sides. One example is the preliminary machine learning work in atrial fibrillation: this has the potential to lead to over diagnosis if there is not enough discussion between the clinical and data science domains. A great fictional comparison is that of Sherlock Holmes (the data scientist) and Dr. Watson (the clinician), the dyad of inspectors that work well together as both neutralize each other's weaknesses. Lastly, the present state of imbroglio in medicine and its future solutions based on artificial intelligence mandates a special duality and synergy of clinicians and data scientists (like the double helices of DNA) with their respective different modes of processing contributing without hubris. Take the integer of "5" and have the clinician and data scientist each be this integer: If the two parties are antagonistic, then it is 5/5 or 1; if the two parties are complementary but not synergistic, then it is 5+5=10; if the two parties are synergistic, then it is 5x5=25.

Small (not always big) data in biomedicine. Finally, for the future, if we have sustained human-to-human collaboration and if we have design deep learning in healthcare and medicine, we will need to utilize not only deep learning but also variants of deep learning for solutions as there cannot be always Big Data to satisfy the ideal situation for deep learning. We need to work with "small" data (individual patients' serial data for example) that is extremely important for many clinicians in serial followup. We also need to think of innovative deep learning methods such as one-shot learning and deep reinforcement learning for less than Big Data.

Making the visible invisible and the invisible visible. The AI projects should render present day paraphernalia such as computers and biometric devices obsolete (making the visible invisible). One such AI-inspired tool is the intelligent agent (or chat bots) that will replace many humans. Conversely, the signal in the noise will be picked up by deep learning and other techniques to make the invisible knowledge and intelligence in healthcare and medicine more discernible. It is good to have someone knowledgeable to triage the situations to see which will be best suited for AI methodologies.

Dual training of clinicians and data scientists. The future clinician may benefit from a dual education in data science and artificial intelligence. These dual-trained scientists can then serve as valuable liaisons between the clinical and data science domains. It is possible that deep learning and AI can render the clinicians less clinically astute so these clinician-data scientists can mitigate this loss of clinical acumen. Lack of a dual perspective can easily lead to false presumptions; overdiagnosis is a potential problem such as studies that have excluded a clinician input. This cohort will also be useful to demystify the “black box” issue of artificial intelligence. The cohort is helpful to police false discovery with finding spurious associations in the training set that have little relevance to new data.

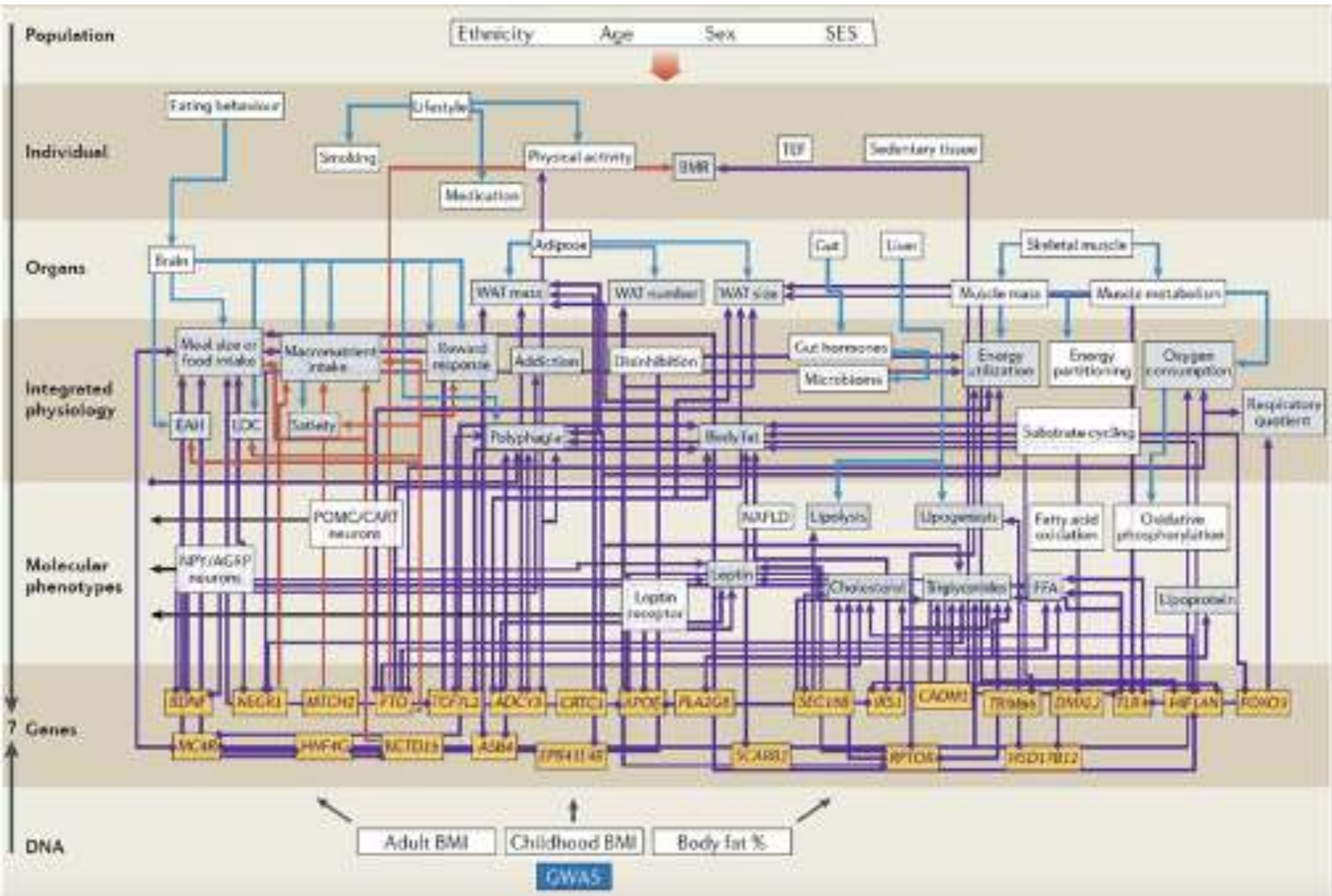
It is not all just about deep learning. This lack of a meaningful continual human (clinician)-to-human (data scientist) interface then results in a paucity of best solutions to real problems in healthcare and medicine with a full understanding of all the nuances. In addition, we have the possibility of creating a myriad of over-



diagnoses which can result in inappropriate therapy and its inherent complications.

In short, we do not, in health care, have “design" learning. It is prudent to implement basic techniques in statistical analysis rather than always thinking that deep learning is the only methodology. Some of the knowledge can be culled from good basic data analytics without heavy duty deep learning and it is important to appreciate when situations do not demand AI. Conversely, many future problems in biomedicine and healthcare will need to be solved with a cognitive solution and not simply deep learning.

Deep phenotyping for complicated nature of biomedicine. The "deep" phenotype of biomedicine will include hundreds of layers of data including genomic and pharmacogenomic profile, socioeconomic milieu, psychological profile, etc. This deep phenotyping necessary to demystify the imbroglio of clinical biomedicine and precision medicine will become far more complex as the years of AI in healthcare and medicine become more omnipresent (see Figure). There is a potential mismatch between precision medicine and population health; AI in biomedicine may be useful to reconcile these two forces.



AI in medicine needs a universal nervous system with a brain. All stakeholders involved in AI in medicine and healthcare should have a rudimentary understanding of the types of artificial intelligence and all their nuances and limitations. Appreciation of the third wave (neuroscience) of AI will be vital for all practitioners of AI in healthcare and medicine. The advent of internet of things (IoT) will evolve into the AI-inspired internet of everything (IoE) in which devices will have some embedded primitive AI capabilities; this is analogous to afferent peripheral nerves connecting to a central nervous system. With this capability, each person's healthcare could be provided a "clinical GPS" with illnesses being "traffic congestions" to be navigated around.

In conclusion, the future of artificial intelligence in medicine is extremely propitious with a myriad of advanced AI techniques such as deep reinforcement learning, one-shot learning, and capsule network that will need to be in synergy with clinicians to allow data to be an enabler of new knowledge and intelligence in biomedicine and healthcare. All healthcare data will need to be liberated and shared without any obstacles so that AI can be ubiquitous and invisible in the future health care arena and discover new knowledge from all sources of data and information. In addition, there needs to be an interface between clinicians with data and computer scientists with analytics to assure a data-to-information continuum and eventually a knowledge-to-intelligence transfer. Finally, we need to promulgate a human-machine synergy via a clinician-data scientist collaboration without hubris to push future healthcare and medicine to the highest echelon.

Finally, what was enlightening about the Human to AlphaGo competition was not only the brilliant and innovative 37th move made by *AlphaGo* during the second game, but also the 78th move of the meaningless fourth game (as AlphaGo had already won three games in the five-game match) in which Lee Sedol made an equally brilliant move. This clearly demonstrated that the computer had a positive influence on man and a man-to-computer synergy would yield the best result for artificial intelligence, especially in medicine and healthcare.

With AI in medicine and healthcare, it is not man versus machine, but man *and* machine.

ARTIFICIAL INTELLIGENCE IN MEDICINE COMPENDIUM

KEY REFERENCES

*	Recommended
**	Highly recommended

Books on Artificial Intelligence

* Armstrong, Stuart. *Smarter Than Us: The Rise of Machine Intelligence.* Machine Intelligence Research Institute, Berkeley, 2014.
- *A short treatise on AI with a well-rounded overview of the relevant current issues.*

Boden, Margaret A. *AI: Its Nature and Future*. Oxford University Press, New York, 2016.

* Bostrom, N. *Superintelligence: Paths, Dangers, Strategies.* Oxford University Press, Oxford, 2014.
- *An outstanding perspective on all the nuances of AI in the current era and in the future.*

Coppin, Ben. *Artificial Intelligence Illuminated.* Jones and Barlett Publishing, Sudbury, MA, 2004.

Christian B and Griffiths T. *Algorithms to Live By: The Computer Science of Human Decisions.* Henry Holt and Company LLC, New York, 2016.

* Domingos, P. *The Master Algorithm: How the Quest for the Ultimate Learning Machine Will Remake Our World.* Basic Books, New York, 2015.
- *A very good book on machine learning for anyone who would like to know more about machine learning beyond the average media publication but would not like to be buried by the esoteric allusions of computer and data science.*

* Hawkins, J. *On Intelligence. How a New Understanding of the Brain will Lead to the Creation of Truly Intelligent Machines.* Times Books, 2004.
- A very good book on how to relate the brain and neuroscience to computers by the founder of Palm Computing.

* Kurzweil, Ray. *How to Create a Mind: The Secret of Human Thought Revealed.* Penguin Books, New York, 2012.
- A fascinating look at the mind and future of artificial intelligence from the futurist Ray Kurzweil.

Lucci S and Kopec D. *Artificial Intelligence in the 21st Century.* Mercury Learning and Information, Dulles, VA, 2013.

** Marcus, G and Freeman J. The Future of the Brain: Essays by the World's Leading Neuroscientists. Princeton University Press, Princeton, NJ, 2015.
- A must read for anyone who would like to stay ahead in understanding the future underpinnings of artificial intelligence as a cognitive science.

** Russell S and Norvig P. *Artificial Intelligene: A Modern Approach* (3rd edition). Pearson Education, Inc, Upper Saddle River, NJ, 2010.
- The most comprehensive and authoritative textbook on artificial intelligence with an incredible historical and futuristic perspective as well as amazing depth and breadth.

* Tegmark M. *Life 3.0: Being Human in the Age of Artificial Intelligence*. Penguin Random House LLC, New York, 2017.
- *A lengthy treatise on AI and its impact on the future of life in all dimensions by an MIT physicist.*

* Wilson JH and Daugherty P. *Human + Machine: Reimagining Work in the Age of AI.* Harvard Business Review Press, Boston, 2018.
- A well-rounded perspective on how AI has changed the dynamics of the workplace by creating a large "middle" or distance that needs to be navigated and connected between man and machine.

Books on Data and Artificial Intelligence in Biomedicine

* Agah A. *Medical Applications of Artificial Intelligence.* CRC Press, Boca Raton, FL, 2014.
- A comprehensive and current textbook on medical applications of AI with more emphasis on the computer and data science aspects.

Campion F and Carlsson G. *Machine Intelligence for Healthcare*.

Cleophas TJ and Zwinderman AH. *Machine Learning in Medicine*. Springer, New York, 2013.

Dua S and Chowriappa P. *Data Mining for Bioinformatics.* CRC Press, Boca Raton, 2014.

Liebowitz J and Dawson A. *Actionable Intelligence in Healthcare.* CRC Press, Boca Raton, FL 2017.

Natarajan P, Frenzel JC, and Smaltz DH. *DemystifyingBig Data and Machine Learning for Healthcare*. CRC Press, Boca Raton, FL 2017.

* Reddy CK and Aggarwal CC. *Healthcare Data Analytics.* CRC Press, Boca Raton, FL 2015.
- The most complete book on data analytics in healthcare with broad range of topics including natural language processing, visual analytics, clinical decision support systems, computer-assisted medical image analysis systems, and information retrieval.

* Szolovits P. *Artificial Intelligence in Medicine.* AAAS Selected Symposia Series, Volume 51. Westview Press Inc, Boulder, Colorado, 1982.
- A very good perspective on the state of the art during this early period of artificial intelligence in medicine with emphasis on expert systems.

Wachter R. *The Digital Doctor: Hope, Hype, and Harm at the Dawn of Medicine's Computer Age.* McGraw Hill, New York, 2015.

Yang H and Lee EK. *Healthcare Analytics: From Data to Knowledge to Healthcare Improvement*. John Wiley and Sons, Hoboken, New Jersey, 2016.

Journals and Journal Issues on Artificial Intelligence

Artificial Intelligence in Medicine (AIM)

Journal of Medical Artificial Intelligence (JMAI)

MIT Technology Review 2017. The Artificial Intelligence Issue.

Articles on Artificial Intelligence and Artificial Intelligence in Medicine

Artificial Intelligence and Life in 2030: One Hundred Year Study on Artificial Intelligence. https://ai100.stanford.edu.

Preparing for the Future of Artificial Intelligence. Executive Office of the President: National Science and Technology Council and Committee on Technology. October, 2016.

Angermueller C et al. Deep Learning for Computational Biology. *Mol Syst Biol* 2016; 12 (878)1-16.

Chang AC. Big Data in Medicine: The Upcoming Artificial Intelligence. *Prog in Ped Card* 2016; 43:91-94.

Coiera EW. Artificial Intelligence in Medicine: The Challenges Ahead. *J of Am Med Informatics Assoc* 1996; 3(6): 363-366.

Deo RC. Machine Learning in Medicine. *Circulation* 2015; 132:1920-1930.

Hanson CW and Marshall BE. Artificial Intelligence Applications in the Intensive Care Unit. *Crit Care Med* 2001; 29:427-435.

Miotto R et al. Deep Patient: An Unsupervised Representation to Predict the Future of Patients From the Electronic Health Records. *Nature Scientific Reports* 2017.

Obermeyer Z and Emanuel EJ. Predicting the Future- Big Data, Machine Learning, and Clinical Medicine. *N Eng J Med* 2016; 375:13-16.

Ramesh AN, Kambhampati C, Monson JR et al. Artificial Intelligence in Medicine. *Annals of the Royal College of Surgeons of England* 2004; 86(5): 334-338.

Ravi D et al. Deep Learning for Health informatics. *IEEE Journal of Biomedical and Health Informatics* 2017; 21(1): 4-21.

Sacchi L and Holmes JH. Progress in Biomedical Knowledge Discovery: A 25-year Retrospective. *Yearb Med Inform* 2016; S117-129.

Quotes

The end of theory: The data deluge makes the scientific method obsolete (Chris Anderson, *Wired*).

AI is the new electricity (Andrew Ng, Baidu)

Artificial intelligence is our biggest existential threat (Elon Musk, Tesla)

The development of full artificial intelligence could spell the end of the human race (Stephen Hawking, Physicist)

Artificial Intelligence is almost a humanities discipline. It is really an attempt to understand human intelligence and cognition (Sebastian Thrun, Computer Scientist)

The question of whether a computer can think is no more interesting than the question of whether a submarine can swim. (Edsger Dijkstra)

(ACC and MI3)

AI in Medicine can Transform a Doctor from Dr. Watson into Sherlock Holmes with His Attention to Data and Intelligence.

The Orchestra of AI in Medicine has Great Musicians (Data Scientists) but We Need the Clinicians to be Composers and Dual-trained Doctors-Data Scientists to be Conductors.

Artificial Intelligence in Medicine, or "Medical Intelligence", Will Make the Visible INVisible and the INVisible Visible.

Medical Data is the Fuel, Artificial Intelligence is the Rocket, and Those Who Use Artificial Intelligence in Medicine and Healthcare Will Be the Astronauts for Moonshot Projects.

We are Rapidly Transitioning from the Former Era of Evidence-Based Medicine to a New Era of "Intelligence-Based Medicine".

Human (Clinician/ Nursing/ Administrator) Intelligence + Machine Intelligence = Medical Intelligence

Websites

AIMed (Ai-Med.io)
Data Science Institute (DSI) at American College of Radiology (ACR)(www.acrdsi.org)
Healthcare.ai

Movies/ VideoClips

Do You Trust This Computer?
A documental focused on the current state of the art on artificial intelligence with emphasis on robots but well rounded. Many luminaries of AI such as Stuart, Andrew Ng, Elon Musk are in this documentary. There are relevant segments in medicine and healthcare and future of AI in these domains.

GLOSSARY

A/

Accountable Care Organization (ACO)/ Creation of entities to accommodate three aims: better care for individuals; better health for populations; and slower growth in costs through improvements in care. (Berwick DM. Launching Accountable Care Organizations- The Proposed Rule for the Medicare Shared Savings Program. *N Engl J Med* 2011; 364: e32.)

Activation Function/ An activation function is a node added to the output end of any neural netowrk.

Adaptive Boosting (or AdBoost)/ A meta-algorithm used in machine learning to improve performance by primarily reducing bias.

Affordable Care Act (ACA)/ The comprehensive health reform signed in March 2010 by President Obama into law to render preventive care more accessible and affordable starting in January 2014. (McDonough JE. The Road Ahead for the Affordable Care Act. *N Engl J Med* 2012; 367: 199-201.)

Agent (see Intelligent Agent)

Algorithms/ Term describing the computer process of following a well-defined list of instructions with historical origin traced to Al Khwarizmi and later popularized by Leonardo Fibonacci. (Steiner C. *Automate This: How Algorithms Came to Rule Our World.* The Penguin Group, New York, 2012.)

AlphaGo/ Google DeepMind group developed this computer program that utilizes a Monte Carlo tree search with an artificial neural network to play the game Go.

American Standard Code for Information Interchange (ASCII)/ A character-encoding scheme to represent common text in computers and related equipment or devices that display text.

API/ (see Application Programming Interface)

Application Programming Interface (API)/ A set of programming instructions and standards (commands, functions, protocols, and objects) for accessing a Web-Based software application or Web tool.

Application Specific Integrated Circuit (ASIC)/ An integrated circuit that is designed for a specific use.

Artificial Conversational Entity (see Chatbot).Artificial General Intelligence (AGI) (see Strong AI).

Artificial Intelligence (AI)/ The science and engineering of making intelligent machines, especially intelligent computer programs (John McCarthy, Stanford). (Chang AC et al. Artificial Intelligence in Pediatric Cardiology: An Innovative Transformation in Patient Care, Clinical Research, and Medical Education. *Congenital Cardiology Today*; 2012.)

Artificial General Intelligence (AGI)/ Also known as strong AI (or true A), this is machine intelligence that can perform any intellectual task done by a human and even at superhuman levels.

Artificial Neural Network (ANN)/ A computational model inspired by natural neurons with communication channels between neurons and these signals that can be weighted (positive or negative). (Lisboa PJ et al. The Use of Artificial Neural Networks in Decision Support in Cancer: A Systematic Review. *Neural Networks* 2006; 19(4): 408-415.)

ASCII (see American Standard Code for Information Interchange).

Association Analysis/ Data mining methodology which is useful for discovering interesting relationships hidden in large data sets that can be represented in the form of association rules (sets of frequent items). (Tan PN et al. *Introduction to Data Mining*, Pearson Education Inc, Boston, 2006.)

Augmented Intelligence (or Intelligence Augmentation)/ Use of technology in a supportive role to enhance human intelligence. This term is sometimes used in place of artificial intelligence.

Augmented Reality (AR)/ 3-D virtual objects are integrated into a 3-D real environment in real-time as a form of advanced computer-assisted navigation or visualization technology. (Ewers R et al. Basic Research and 12 Years of Clinical Experience in Computer-assisted Navigation Technology: A Review. *Int J Oral and Max Surg* 2005; 34(1): 1-8.)

Autoencoder/ A type of artificial neural network that is used in unsupervised learning algorithm in order to achieve backpropagation.

Automation Bias/ A group of errors that humans have a tendency to make in highly automated decision-making situations. Elements of automation bias include: errors of omission and commision, overreliance, and overcompliance.

Avatar/ A graphical representation of oneself in the virtual world. (Hansen MM. Versatile, Immersive, Creative and Dynamic Virtual 3-D Healthcare Learning Environments: A Review of the Literature. *J Med Internet Res* 2008; 10(3): 226.)

B/

Backpropagation/ A fast algorithm used for learning in computing gradients for artificial neural networks.

Backward Chaining (or Reasoning)/ An inference method that is described as working backward from the goal to determine if there is data to support these outcomes.

Bag of Words (BoW)/ A model to represent text data in machine learning that involves disregarding the order or structure of words in the document.

Bayesian (or Bayes) Network (BN)(also Belief Networks or Bayes Nets)/ A graph-based model that encodes the probabilistic relationships or dependencies among the variables of interest, thereby can be used to learn causal relationships. (Friedman N et al. Using Bayesian Networks to Analyze Expression Data. *J Computational Biol* 2004; 7(3-4): 601-620.)

Belief Network/ see Bayesian Network.

Big Data/ Recent paradigm describing the coupling of the massive amount of data with sophisticated data analytics to acquire new knowledge or insight. (Chang AC. Big Data in Pediatric Cardiology: The Upcoming Knowledge Revolution. *Congenital Cardiology Today* 2013; 11:11-12).

Biocybernetics (also Biocybernetic System)/ A method to evaluate the entire living organism with an interdisciplinary approach that involves laboratory data using biomodeling and computational elements.

Bioinformatics/ An interdisciplinary study of storage, retrieval, organization, and analysis of biological data and information that utilizes computer science, biology, and mathematics. (Bayat A. Bioinformatics. *Br Med J* 2002; 324(7344): 1018-1022.)

Biomedical Informatics/ An interdisciplinary field of quantitative and computational methods designed to solve problems across the entire spectrum from biology to medicine. (Altman R. Challenges for Biomedical Informatics and Pharmacogenomics. *Ann Rev Pharmacol and Toxicol* 2002; 42: 113-133.)

Biomedical Signal Analysis/ A visualization and interpretation method in biology and medicine for detection, storage, transmission, analysis, and display of images.

Biometrics/ Use of biological markers (such as genetic information, fingerprints, voice patterns, facial characteristics, or eyes) for identification or quantification purposes with the aid of technology.

Biomimicry/ An innovative methodology that is inspired by nature to seek a sustainable and ecological solution. (Sommer AP et al. Biomimicry Patterning with Nanosphere Suspensions. *Nano Letters* 2003; 3(5): 573-575.)

Bionics/ Body parts that are made stronger or more capable by special electronic or electromechanical devices; use of a biological entity and electronic devices that are put together to create the required implant. (Wallace GG et al. Organic Bionics: Molecules, Materials, and Medical Devices. *Chemistry in Australia* 2009; 76(5): 3-8.)

Blockchain (or Block Chain)/ An innovative secure database used in bitcoin that is shared by all parties in a distributed network of computers so that an anonymous exchange of digital assets can occur.

Blog/ From “web” and “log”, a virtual informational site that consist of users placing posts (or entries). (Van De Belt TH et al. Definition of Health 2.0 and Medicine 2.0: A Systematic Review. *J Med Internet Res* 2010; 12(2): e18.)

Boolean Algebra/ A branch of algebra in which the values of the variables are the truth values true and false (denoted 1 and 0, respectively).

Bots (see Chatterbot).

Brain Based Device (BBD)/ Synthetic neural model with behavioral tasks embodied in robotic phenotypes, or a neurally controlled robot.

Brain Computer Interface (BCI)/ Also called mind-machine interface (MMI) or brain-machine interface (BMI), is a communication between the brain and computer to enhance or augment either mental or physical ability. (Lee B et al. A Primer on Brain-Machine Interfaces, Concepts, and Technology: A Key Element in the Future of Functional Neurorestoration. *World Neurosurg* 203; 79(3-4): 457-471.)

C/

Capsule Network (also Capsule or CapsNet)/ A relatively new deep learning concept promulgated by Geoff Hinton that is based on human brain modules called "capsules" that are good for routing visual images to the appropriate capsule (Sabour S et al. Dynamic Routing Between Capsules. 31st Conference on Neural Information Processing Systems (NIPS 2017).

Case-based Reasoning/ Artificial intelligence technique of utilization of former experiences to comprehend and solve new problems. (Holt A et al. Medical Applications in Case-based Reasoning. *The Knowledge Engineer Rev* 2005; 20(3): 289-292.)

Central Processing Unit (CPU)/ The electronic circuitry in the computer that performs the basic functions and operations (such as arithmetic, logical, control, and input/output) specified by the instructions.

Chaos Theory/ The study of nonlinear dynamics in mathematics in which seemingly random events are predictable from simple deterministic equations. (Holm S. Does Chaos Theory Have Major Implications for Philosophy of Medicine? *Med Humanities* 2002; 28: 78-81.)

Chatbot (also Chatterbot or Bot)/ A software application that performs certain automated tasks with web spidering being one of the largest uses of this modality.

Chimeraplasty/ A gene-therapy technique that involves using synthetic strand of RNA and DNA to form a chimeraplast, which in turn attaches to a target gene. (Taubes G. The Strange Case of Chimeraplasty. *Science* 2002; 298:2116-2120.)

Classification Trees/ see Decision Trees

Clinical Decision Support System (CDSS)/ An information system that entails tools (such as computerized alerts, clinical guidelines, and condition-specific order sets) that provides the clinical staff data and information that can improve health care. (Garg AX et al. Effects of Computerized Clinical Decision Support on Practitioner Performance and Patient Outcomes: A Systematic Review. *JAMA* 2005; 293(10): 122—1238.)

Circos Plot/ A visualization tool using circular ideogram layout to identify similarities and differences from genomes. (Krzywinski M et al. Circos: An Information Aesthetic for Comparative Genomics. *Genome Res* 2009; 19: 1639-1645.)

Classifier or Classification/ An algorithm that implements a classification in machine learning such as naive Bayes, random forest, boosted trees, or support vector machine.

Clinical Document Architecture (CDA)/ A markup standard from HL7 that defines the medical record structures for certain documents like discharge summaries and progress notes to facilitate exchange of information between providers and patients.

Cloning/ Creation of an organism that is an exact genetic copy of another with every DNA material being identical. (McLaren A. Cloning: Pathway to a Pluripotent Future. *Science* 2000; 288(5472): 1775-1780.)

Cloud Computing (also Cloud)/ A jargon term to describe the workload shift via the internet from local computers to a remote network of computers (infrastructure, platform, and software as a service, or IaaS, PaaS, and SaaS, are models). (Soman AK. *Cloud-based Solutions for Healthcare IT*. CRC Press, New York, 2011.)

Cluster Analysis/ A data science unsupervised machine learning and data science methodology of dividing data or objects into groups or clusters that exude the same or similar characteristics.

CNTK (Microsoft)/ Cognitive Toolkit is an open source deep learning framework.

Codon/ A three-nucleotide sequence of DNA or RNA that specifies a single amino acid. (Hudson KL. Review Article: Genomics, Health Care, and Society. *N Engl J Med* 2011; 365: 1033-1041.)

Cognitive Computing/ System (such as Watson from IBM) that can learn better than man or computer alone with the intersection of three disciplines: neuroscience, supercomputing, and nanotechnology. (Brasil LM et al. Hybrid Expert System for Decision Supporting in the Medical Area: Complexity and Cognitive Computing. *Int J of Medical Informatics* 2001; 63(1-2): 19-30.)

Combinatorics (or Combinatorial Mathematics)/ Field of mathematics that is concerned with problems of selection, arrangement, and operation within a fine or discrete system.

Comma-Separated Value (CSV)/ A file that stores data in tabular structured format and have the extension .csv.

Complementary Learning/ An AI system that explores both deep learning and cognitive memory to work in unison so that one-shot learning using associative and episodic memories can be sed for individual dynamic patterns.

Complementary-DNA (c-DNA)/ A DNA molecule synthesized by an RNA-dependent DNA polymerase from an RNA template. (Hudson KL. Review Article: Genomics, Health Care, and Society. *N Engl J Med* 2011; 365: 1033-1041.)

Complexity Theory (see Chaos Theory)

Computer-Aided Detection (CADe)/ A technology involved in decreasing false negative detection rates of physicians interpreting the images.

Computer-Aided Diagnosis (CADx)/ A technology

Computer Assisted Design (CAD)(also Computer Aided Design and Drafting, or CADD)/ The utilization of computer systems and software to aid the creation, modification, and optimization as well as storage of a particular design. (Grassy G et al. Computer-assisted Rational Design of Immunosuppressive Compounds. *Nature Biotechnol* 1998; 16: 748-752.)

Computer-Generated Imagery (CGI)/ Three-dimensional computer graphics used for creating scenes or special effects. (McCloy R. Virtual Reality in Surgery. *BMJ* 2001; 323(7318: 912-915.)

Computer Vision/ see Machine Vision.

Concatenation/ The linkage of two or more character strings (words) or separated things end to end to be addressed as a single item (e.g. air and line can be concatenated into airline).

Continuity of Care Document (CCD)/ An XML-based markup standard designed for a patient summary clinical document that can be shared by all computer applications.

Convolution/ A mathematical function derived from two given functions by integration that expresses how the shape of one is modified by the other (the blending of one function with another).

Convolutional Neural Network (CNN)/ A special type of forward-feeding ANN used for image detection and interpretation in which the neurons are inspired by the cortex of the brain with overlapping influence.

Copy Number Variation (CNV)/ Variation from one person to the next in the number of copies of a particular gene or DNA sequence. (Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med* 2010; 363: 166-176.)

CPU (see Central Processing Unit).

Crowd-sourcing/ Distributive problem solving for services or ideas by distributing tasks to a large on-line community to mine collective intelligence. (Swan M. Health 2050: The Realization of Personalized Medicine Through Crowdsourcing, the Quantified Self, and the Participatory Biocitizen. *J Pers Med* 2012; 2: 93-118.)

CURES Act (21st Century CURES Act)/ Legislation, passed in 2016, designed to accelerate medical product development (drugs and devices) in order to advance innovations.

Curse of Dimensionality/ How certain learning algorithms in machine learning do not perform well in high-dimensional data.

Cybersecurity/ The technologies and practices that are designed for the protection of computer systems and information from theft and disclosure.

Cyborg/ From “cybernetic” and “organism, a being with biological and artificial parts. (Williams SJ. Modern Medicine and the ”Uncertain Body”: From Corporeality to Hyperreality? *Social Science and Medicine* 1997; 45(7): 1041-1049.)

Cyc/ An AI project started by Stanford computer scientist Douglas Lenat that focused on using human-like reasoning in collecting ontology and knowledge base concepts; considered by some the precursor to IBM Watson.

D/

Data Analytics/ Process of transforming and modeling data for discovery useful information.

Data Dictionary/ A central repository of information about data in a data model for programmers and others who need the information.

Data Exhaust/ Data generated as information byproducts resulting from all of digital as well as online activities and consist of storable choices, actions and preferences.

Data Frame/ A storage mechanism for data tables in programming language R (e.g. df=data.frame(a,b,c) where a,b,and c are vectors).

Data Hub/ A collection of data from a variety of sources organized for sharing or distributing that is different than a data warehouse (unintegrated) or a data lake (data homogenized).

Data Lake/ A storage mechanism for all data within a system in its natural format; usually favored by data scientists (as opposed to data warehouses being favored by business intelligence experts).

Data Mart/ A subset of data warehouse that focus on a particular subject or department. (Arnrich B et al. Data Mart based Research in Heart Surgery: Challenges and Benefit. *Medinfo* 2004; 11 Pt 1; 8-12.)

Data Mining/ The process of automatically discovering useful information in large data repositories. (Tan PN et al. *Introduction to Data Mining*, Pearson Education Inc, Boston, 2006.)

Data Reservoir/ A managed and organized version of data lake in which there are data platforms (such as Apache Hadoop and NoSQL databases as well as relational database servers) to serve both the data scientist as well as the business users.

Data Science/ Interdisciplinary field (statistics, machine learning, data mining, predictive analytics, and mathematics) about strategies to extract knowledge from data in structured and unstructured forms.

Data Visualization/ Discipline to study the visual representation of data to maximize communication with clarity. (Chittaro L. Information Visualization and its Application in Medicine. *Artif Intel in Med* 2001; 22(2): 81-88.)

Data Warehouse (DW)/ A database or collection of databases for reporting and data analysis to be used for management decision making. (Wisniewski MF. Development of Clinical Data Warehouse for Hospital Infection Control. *J Am Med Inform Assoc* 2003; 10(5): 454-462.)

Data-Driven Medicine/ Data-centric approach that compute on massive amounts of data to discover previously unrecognized patterns and to make clinical relevant predictions to improve health care. (Shah NH et al. The Coming Age of Data-driven Medicine: Translational Bioinformatics' Next Frontier. *J Am Med Inform Assoc* 2012; 19:e2-e4.)

Datathon/ A venue for data scientists (and other disciplines depending on the theme) to work closely and intensely together similar to hackathons.

Decision Theory/ Identification of value and uncertainty relevant to a decision and process of determining the optimal solution.

Decision Trees/ Decision support mechanism to display an algorithm that uses a tree-figured graph or model of decisions with consequences.

Deductive Reasoning/ Starting with general theory or observation that lead to data (top down).

DeepMind (Google)/ London-based computer scientist and neuroscientist group headed by Demis Hassabis that developed the AlphaGo software that defeated the human champion.

DeepQA Project/ A project aimed at illustrating how the advancement and integration of natural language processing, information retrieval, machine learning, and knowledge representation to accommodate open domain questions as observed with the IBM Watson supercomputer. (Ferrucci D. Build Watson: An Overview of DeepQA for the Jeopardy! Challenge. Proceeding of the 19th International Conference on Parallel Architectures and Compilation Techniques 2010; 1-2.)

Deep Neural Networks/ Computationally powerful version of regular neural networks with two or more layers of hidden processing neurons.

Deep Learning/ An artificial intelligence technique that extends the traditional machine learning techniques with multiple layers of neural networks.

Defense Advanced Research Projects Agency (DARPA)/ A federal agency, established in 1958, to prevent strategic surprise from negatively impacting U.S. national security by maintaining the technological superiority of the U.S. military in many disciplines; the innovative engine for the Department of Defense. (Polla DL et al. Microdevices in Medicine. *Ann Rev fo Biomed Engineer* 2000; 2: 551-576.)

Dendral Expert System/ An early AI project at Stanford in which an heuristic system is designed to be a chemical analysis expert system.

Dendrimer/ A molecule or polymer with a symmetric repetitively branching three-dimensional pattern; also called a cascade molecule. (Lee CC et al. Designing Dendrimers for Biological Applications. *Nature Biotechnol* 2005; 23:1517-1526.)

DGX-1/ An integrated system for deep learning by NVIDIA that features eight Tesla P100 accelerators connected though NVLink, the NVIDIA high performance GPY interconnect.

Differential Privacy/ The cryptographical process of maximizing accuracy of queries from statistical databases while minimizing chances of identifying its records.

Digital Imaging and Communications in Medicine (DICOM)/ An international standard for storing, displaying, processing, and transmitting medical images for integration of medical image devices so that images are interoperable.

Digital Medicine/ Use of digital tools in medicine to record clinical data and generate medical knowledge that is more precise, more effective, more experimental, and more widely distributed. (Shaffer DW et al. What is Digital Medicine? *Stud Health Technol Inform* 2002; 80:195-204.)

DNA Computing (also Molecular Computing)/ Future computing technology to incorporate DNA and RNA at high and low concentrations to send signals through a computational configuration. (Maojo V et al. Nanoinformatics and DNA-Based Computing: Catalyzing Nanomedicine. *Ped Res* 2010; 67: 481-489.)

DNA Microarray/ Also called DNA chip, it is a system in which many probes with known identity are fixed on a solid support with spots that can be DNA, cDNA, or oligonucleotides. (Dugoff L. Application of Genomic Technology in Prenatal Diagnosis. *N Engl J Med* 2012; 367(23): 2249-2251.)

Drone (also Unmanned Aerial Vehicle (UAV)/ A pilotless air vehicle that is operated by remote control.

Drug Discovery/ Process in which novel candidate medications for diseases are discovered though computational means.

E/

eHealth (also e-health)/ Healthcare accompanied by electronic processes to enhance the communication and quality of delivery. (Black AD et al. The Impact of eHealth on the Quality and Safety of Health Care: A Systematic Overview. *PLoS Med* 8(1): e1000387.)

Electronic Health or Medical Record (EHR or EMR)/ Digital record of a patient's paper health (comprehensive) or medical (narrower) status that is available for authorized users for reading and decision making.

EMR Adoption Model (EMRAM)/ The HIMSS effort to track health institutions and adoption of EMR in stages (0-7) with stage 7 adoption being no use of paper for medical records.

Embedded AI (eAI)/ The application of AI in small devices to push machine and deep learning tools away from only central sources.

Embryonic Stem Cell/ These stem cells are pluripotent stem cells that originate from the blastocyst inner cell mass after an *in vitro* fertilization. (Bishop AE et al. Embryonic Stem Cells. *J Pathol* 2002; 197:424-429.)

Encyclopedia of DNA Elements (ENCODE)/ An international collaboration of research groups funded by the National Human Genome Research Institute (NHGRI) to identify all functional elements in the human genome. (The ENCODE Project Consortium. Identification and Analysis of Functional Elements in 1% of the Human Genome by the ENCODE Pilot Project. *Nature* 2007; 447: 799-816.)

Enigma Machine/ A cipher machine invented by German engineer Arthur Scherbius at the end of World War I to be used for encipher/decipher messages.

Enterprise Resource Planning (ERP)/ A software that integrates various functions of an organization such as CRM, human resources, accounting, and inventory and order management with a shared central database that can support these functions.

Enterprise Data Warehouse (EDW)/ (see Data Warehouse)

Entity Relation (ER) Model or Diagram/ A graphical representation of the entities and the relationships to illustrate data and database infrastructure.

Epigenetics/ The study of the interactions of chemicals and genes and the factors that influence these interactions such as DNA methylation, histone modification, and nucleosome location. (Hochberg Z et al. Child Health, Developmental Plasticity, and Epigenetic Programming. *Endocrine Reviews* 2011; 32(2): 159-224)

Evidence-Based Medicine/ An interdisciplinary strategy to learn by making decisions on research studies that are carefully selected; the five step process includes ask-acquire-appraise-apply-analyze. (Sackett DL et al. Evidence based Medicine: What it is and What it isn't. *BMJ* 1996; 312(7023): 71-72.)

Evolutionary Algorithms (see Genetic Algorithms).

Exabyte (EB)/ One quintillion or 10^{18} bytes (or 1 billion gigabytes) which is enough storage capacity to store 100,000 times all printed material.

Exome/ The part of the genome that contains the DNA record for the protein coding part of the genome. (Ng SB et al. Exome Sequencing Identifies the Cause of a Mendelian Disorder. *Nature Genetics* 2010; 42: 30-35.)

Exon/ A sequence of DNA that codes information for protein synthesis that is transcribed to messenger RNA. (Keren H et al. Alternative Splicing and Evolution: Diversification, Exon Definition, and Function. *Nature Reviews Genetics* 2010; 11:345-355.)

Exoskeleton (also Powered Exoskeleton)/ A powered suit to assist the wearer to increase strength and/or endurance with healthcare uses including lifting of heavy patients and allowing disabled patients to walk. (Hesse S et al. Upper and Lower Extremity Robotic Devices for Rehabilitation and for Studying Motor Control. *Eur J Phys Rehabil Med* 2012; 48(1): 111-121.)

Expert Systems/ An artificial intelligence methodology, among the first successful techniques used in artificial intelligence, in which a computer system emulates the decision-making process of a human expert. (Spiegelhalter DJ et al. Assessment, Criticism, and Improvement of Imprecise Subjective Probabilities for a Medical Expert System. *Proceedings of the Fifth Conference on Uncertainty in Artificial Intelligence* 2013.)

Extract, Transform and Load (ETL)/ A process that occurs in a data repository in which data is extracted out of the source and organized into a data warehouse or other data structure.

Extensible Markup Language (XML)/ A markup language that defines a set of rules for encoding documents for storage and transport in a human- and machine-readable format.

F/

Facebook/ An online social networking site with potential problems including exposure to inappropriate materials, cyberbullying, sexting, and even Facebook depression. (O'Keefee GS et al. The Impact of Social Media on Children, Adolescents, and Families. *Pediatrics* 2011; 127(4): 800-804.)

Fast Healthcare Interoperability Resources (FHIR)(pronounced "fire")/ An interoperability standard for electronic exchange of healthcare information developed by HL7 to provide a framework for exchange to support clinical practice.

Feature/ An individual measurable element of a phenomenon being observed (number, string, or graph) for use in machine learning algorithm for pattern recognition, classification, and regression.

FHIR (see Fast Healthcare Interoperability Resources)

Field Programmable Gate Array (FPGA)/ These are integrated semiconductor devices that utilizes the concept of configurable logic blocks so that these are programmable after manufacturing.

File Transfer Protocol (FTP)/ A standard Internet client-server protocol designed to transmit files between computers over TCP/IP connections.

FLoating point Operations Per Second (FLOPS)/ A computer performance measurement benchmark for rating the speed of microprocessors. For example, one teraFLOPS is equal to one trillion FLOPS.

Forward Chaining/ A process in expert systems in which inference rules are used to extract more data until goal is reached.

Foursquare/ A social networking website designed for mobile devices in order to interact with their environment.(Merchant RM et al. Integrating Social Media into Emergency-Preparedness Efforts. *N Engl J Med* 2011; 365: 289-291).

Fullerenes/ Composed entirely of carbon, these molecules used in nanotechnology can take the shape of spheres, ellipsoids, or tubes. (Raffa V et al. Progress in nanotechnology for Healthcare. *Minim Invasive Ther Allied Technol* 2010; 19(3): 127-135)

Fuzzy Logic or Reasoning/ It is a problem-solving control system that accommodates degrees of truth rather than absolute true or false Boolean logic. (LaBrunda M et al. Fuzzy Logic in Medicine. *J of Inform Technol Res* 2008: 1(1): 1-7.)

G/

Gamification/ The application of digital game design techniques and leveraging of technology and psychology to non-game contexts such as social impact or health care challenges. (Ferguson B. The Emergence of Games for Health. *Games for Health Journal* 2012; 1: 1-3.)

Gene Therapy/ Techniques that involve the use of genes to treat or prevent disease: replacing or inactivating a mutated gene or introducing a new gene. (Kohn DB et al. Gene Therapy Fulfilling its Promise. *N Engl J Med* 2009; 360(5): 518-521.)

General AI (see Artificial General Intelligence)

Generative Adversarial Network (GAN)/ Neural networks that are composed of two separate deep neural networks (called the generator and the other, the discriminator) that are in essence competing each other.

Genetic Algorithms (GA)/ An artificial intelligence technique that mimics biological evolution by representing the solution to the problem as a genome and applying genetic operators to evolve the eventual best solution. (Dybowski R et al. Prediction of Outcome in Critically Ill Patients Using Artificial Neural Network Synthesized by Genetic Algorithm. *The Lancet* 1996; 347(9009): 1146-1150.)

Genetic Engineering/ Technologies that can modify the genetic makeup of cells and can involve highly sophisticated manipulations of genetic material, also called genetic modification. (Gaj T et al. Expanding the Scope of Site-Specific Recombinases for Genetic and Metabolic Engineering. *Biotechnol Bioeng* 2013; [Epub ahead of print].)

Genetic Information Nondiscrimination Act (GINA)/ Signed by President Bush in 2008, the act protects Americans against discrimination based on their genetic information in situations that involve health insurance and job employment. (Korobkin R et al. The Genetic Information Nondiscriminatory Act- A Half-Step Toward Risk Sharing. *N Engl J Med* 2008; 359: 335-337.)

Genetically Modified Organism (GMO)/ A plant or animal that has been genetically engineered with DNA from a source such as bacteria, virus, or another plant or animal. (Frey J. Biological Safety Concepts of Genetically Modified Live Bacterial Vaccines. *Vaccine* 2007; 25(30): 5598-5605.)

Genetics, Robotics, Internet, and Nanotechnology (GRIN) Technologies/ Emerging technologies that follow a curve of exponential change. Another term is Nanotechnology, Biotechnology, Information technology, and Cognitive science (NBIC). (Mulhall G. *Our Molecular Future: How Nanotechnology, Robotics, Genetics, and Artificial Intelligence Will Transform Our World.* Prometheus Books, New York, 2002.)

Genomics/ Genomics is the study of functions and interactions of all the genes in the genome, not just of single genes as in genetics. (Guttmacher AE et al. Review Article: Genomic Medicine. *N Engl J Med* 2002; 347: 1512-1520.)

Genome-wide Association Study (GWAS)/ An approach used in genetics research to look for associations between many (typically hundreds of thousands) specific genetic variations (most commonly single-nucleotide polymorphisms) and particular diseases. (Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med* 2010; 363: 166-176.)

Global Innovation Index (GII)/ A valuable benchmarking tool of innovation based on national economy pillars (such as institutions, human capital and research, infrastructure, market sophistication, and business sophistication) as well as innovation output pillars: knowledge and technology and creative outputs. (http://www.globalinnovationindex.org)

Good Old Fashioned AI (GOFAI)/ Also known as symbolism, GOFAI was the first wave of AI that tried to describe intelligence in symbolic terms.

Google/ A search engine company with the mission "to organize the world's information and make it universally accessible and useful" by using "Googlebots" to crawl and search via algorithms (see above); now with one billion search requests a day. (Tang H et al. Googling for a Diagnosis- Use of Google as a Diagnostic Aid: Internet Based Study. *BMJ* 2006; 333:1143-1145.)

Googlebots/ The search algorithm bot (also called a "spider") that crawls and collects documents from the web for the Google search engine. (Giustini D. How Google is Changing Medicine. *BMJ* 2005; 331(7531): 1487-1488.)

Google Glass/ A wearable personal computer with optical head-mounted display with capability to communicate with the internet with natural language commands with potential for application in medicine and surgery. (www.google.com/glass)

GPU (see Graphical Processing Unit).

Graph Database or Theory/ A NoSQL type of database that uses graph structures for semantic queries with nodes and edges to represent and store data.

Graphene/ A material that is composed of a one-atom-thick layer of carbon that is purported to be much stronger than steel while able to conduct electricity; potential applications include screens and displays, biomedical devices and sensors, and memory chips and electronic processors. (Yao J et al. Chemistry, Physics, and Biology of Graphene-based Nanomaterials: New Horizons for Sensing, Imaging, and Medicine. *J Mater Chem* 2012; 22: 14313-14329.)

Graphical User Interface (GUI)/ An interface that allows the user to be able to communicate with the electronic device via graphical icons and visual indicators.

Graphical Processing Unit (GPU)/ A single-chip processor with highly parallel structure used to manage and boost the performance of video and graphics as well as algorithms for large blocks of data in order to lessen the burden of the CPU. Also known as visual processing unit (VPU).

Ground Truth/ A term to describe a reality check or checking for accuracy for machine learning algorithms against the real world.

H/

Hackathon/ An event in which computer scientists gather to collaborate on coding in a very short time (usually days or weekends).

Hadoop (Apache)/ Software library that serves as a framework for distributed processing of large data sets across clusters of computers using relatively simple programming models; emerging as the leading technology for big institutions to mine big data. (O'Driscoll A et al. 'Big Data', Hadoop, and Cloud Computing in Genomics. *J of Biomed Inform* 2013; 46(5): 774-781.)

Haplotype/ A set of DNA variations or polymorphisms that tend to be inherited together. (Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med* 2010; 363: 166-176.)

Haptic Technology (also Haptics)/ The science of understanding and improving human interaction with the physical world through the sense of touch. (http://haptics.seas.upenn.edu.)

Health 2.0 (also Medicine 2.0)/ Combination of health care with the concept of Web 2.0 with an additional assumption of patient empowerment. (Van De Belt TH et al. Definition of Health 2.0 and Medicine 2.0: A Systematic Review. *J Med Internet Res* 2010; 12(2): e18.)

Healthcare Information and Management Systems Society (HIMSS)/ An international not-for-profit organization focused on better health though information and technology. The HIMSS EMR Adoption Model (EMRAM) tracks health institutions and adoption of EMR in stages (0-7).

Health Insurance Portability and Accountability Act (HIPPA)/ The 1996 act that was endorsed by the U.S. Congress that provided regulations for use and disclosure of an individual's health information. It is now often discussed in the context of data security.

Health Level 7 (HL7)/ A not-for-profit standards developing organization that is dedicated to provide a comprehensive framework for the exchange and integration of electronic health information.

Healthcare Effectiveness Data and Information Set (HEDIS)/ An information tool used by health plans in order to measure 81 care and service performance metrics in 5 domains.

Hebbian Theory or Learning/ A theory of how neuronal connections can be enforced (the strength of the connection altered) and is the basis for weight selection in artificial neural networks.

Heuristic Methods/ A technique that is designed for solving a problem more quickly when traditional methods are too slow.

Hidden Markov Model (HMM)(see Markov Model)/ A popular statistical methodology for modeling a wise range of time series data as well as signal processing, particularly speech processing.

HL7 (see Health Level 7)

Holography (also Hologram)/ 3-D image of an object that is projected and captured on a 2-D surface formed by a split laser beam. (Mirza K et al. Holography in Clinical Anatomy Education: A Systematic Review. *Medical Posters* 2013; 1(4).)

Human-Computer Interaction (HCI)/ The interdisciplinary field of study that involves interaction between human users and computers; computer science, behavioral science, design science, cognitive psychology, and communication theory are all involved in this field. (Greatbatch D et al. Interpersonal Communication and Human-Computer Interaction: An Examination of the Use of Computers in Medical Consultations. *Interacting with Computers* 1993; 5(2): 193-216.)

Human Genome Project/ The quest to sequence all 3 billion base pairs of the human genome that was led by the National Institute of Health and completed in 2003 with discovery of more than 1,800 disease genes. (Lander ES. Initial Impact of the Sequencing of the Human Genome. *Nature* 2011; 470(7333): 187-197.

Hybrid Assistive Limb (HAL)/ A cyborg-type robot with a voluntary control system designed to expand or improve the user's physical capability by receiving biosignals from the skin. (Hong YW et al. Lower Extremity Exoskeleton: Review and Challenges Surrounding the Technology and its Role in Rehabilitation of Lower Limbs. *Aust J of Basic and Appl Sci* 2013; 7(7): 520-524.)

Hypergraphs/ A special kind of graph structure in which an edge (called hyperedges) can join a any number of vertices.

Hypertext Markup Langauge (HTML)/ The standard markup language for creating Web pages.

I/

ImageNet/ Large visual database with over 10 million images for research in visual recognition that spawned the recent AI research.

Imputation/ The process of replacing missing data with substituted values.

Inductive Reasoning/ Starting with data and ending with general observation (bottom up).

Inductive Logic Programming/ This is the intersection between machine learning and logic programming and can lead to automated learning of logic rules from eamples and background knowledge.

Inference Engine/ A component of the expert system along with knowledge base with its stored facts that applied logical rules to the knowledge base.

Informatics for Integrating Biology and the Bedside (i2b2)/ An NIH-funded center for developing a scalable informatics framework that will enable clinical research with the existing clinical data for discovery research.

Information Communications Technology (ICT)/ The infrastructures and components needed for modern computing: cloud computing, software, hardware, transactions, communications technology, data, and internet access.

Information Technology (IT)/ The use of computers and other devices to create and store as well as secure all forms of electronic data.

In-Frame Exon Skipping/ The skipping of an exon that contains a multiple of three nucleotides during splicing of pre-mRNA, resulting in the preservation of the reading frame for translation. (Feero WG et al. Review Article: New Therapeutic Approaches to Mendelian Disorders. *N Engl J Med* 2010; 363: 852-863.)

Innovation/ Defined as the act of introducing something new or different that creates value. (Herzlinger RE. Why Innovation in Health Care is So Hard. *Harvard Business Review* 2006.)

Intelligence-Based Medicine (ACC)/ Medicine that is based on the aggregate of evidence-based medicine, data from all sources, and knowledge that is not published from caretakers.

Intelligence Amplification (see Augmented Intelligence). Intelligence Augmentation (see Augmented Intelligence).

Intelligent Agent (also Agent)/ A unit that can perceive its environment via sensors and act upon the environment via effectors. (Vicari RM et al. A Multi-agent Intelligent Environment for Medical Knowledge. *Art Intel in Med* 2003; 27: 335-366.)

Intelligent Clinic (iClinic) (ACC)/ A futuristic clinic in which technology such as wearable devices, intelligent agents, machine learning, and telepresence are utilized to provide a reassuring home environment for patients.

Intelligence Computing Method (ICM)/ Artificial intelligence methods that include genetic algorithm, artificial neural network, and fuzzy logic. (Pandey B et al. Knowledge and Intelligent Computing System in Medicine. *Computers in Biol and Med* 2009; 39(3): 215-230.)

Interaction Networks/ A model which can reason about how objects in a complex system interact so that dynamic prediction can be made.

Internet Protocol (IP)/ Protocol by which data is send from one conputer to another on the Internet. Each computer has a unique IP address.

Internet of Everything (IoE)/ The intelligent connection of people, process, data, and things to make networked connections from IoT (see below) more relevant and valuable.

Internet of Things (IoT)/ The infrastructure of the information society that links together networks, devices, and data for the purpose of collecting and exchanging data.

Interoperability/ The extent that systems and devices can exchange and share data and interpretation.

Intron/ A noncoding segment of DNA between exons that can interrupt a gene-coding sequence. (Guttmacher AE et al. Review Article: Genomic Medicine. *N Engl J Med* 2002; 347: 1512-1520.)

IP/ (see Internet Protocol)

IRIS Platform (Roche)/ Digitial pathology solution that captures all the steps of the histopathological workflow.

Isabel/ A Web-based diagnostic checklist and decision support tool to help clinicians broaden their differential diagnosis and recognize a disease at the point of care (www.isabelhealthcare.com).

J/

JavaScript Object Notation (JSON)/ An open-standard data format that is used for asynchronous browser/server communication with human-readable text to transmit data objects.

Julia/ A high-level and high-performance programming language that is designed to enable high-performance numerical analyses and computational science.

Jupyter (or Jupyter Notebooks)/ An open source project (named after three common programming languages: Julia, Python, and R) rooted in Python (as IPython Notebook) that is used by data scientists and engineers as a tool for collaboration.

K/

Kaggle/ A crowdsourcing approach to a platform for predictive modeling and analytics competitions for the best predictive model that is transparent.

Keras/ A high-level neural network API in Python for fast experimentation.

k-Means Clustering/ A signal processing method of vector quantization for cluster analysis in data mining by which a dataset is partitioned into small number of clusters (by minimizing the distance between each data point and the center of the cluster that the point belongs to).

k-Nearest Neighbor (KNN)/ A machine learning algorithm that is a non-parametric method for classification and regression for pattern recognition.

Knowledge-Based Systems (KBS)/ Artificial intelligence techniques that include rule-based learning, case-based learning, and model-based learning; the other school is intelligent computing method. (Pandey B et al. Knowledge and Intelligent Computing System in Medicine. *Computers in Biol and Med* 2009; 39(3): 215-230.)

Knowledge Discovery in Databases (KDD)/ Data mining for searching hidden knowledge in a large amount of data with data preparation, selection, cleansing, and intrepretation.

Knowledge Representation/ An artificial intelligence technique to represent knowledge in symbols in order to derive at conclusions from these elements. Peleg M et al. Decision Support, Knowledge Representation, and Management in Medicine. *Yearb Med Inform* 2006; 45: 72-80.)

Kolmogorov-Smirnov Test/ This is a non-parametric statistics test that comparesa known hypothetical probability distribution to teh distribution generated by one's data.

L/

Lab-on-a-Chip (also Microfluidics)/ Performing laboratory operations on a small scale (to tens of micrometers) using miniaturized microfluidic devices. (Whitesides GM. Overview: The Origins and the Future of Microfluidics. *Nature* 2006; 442: 368-373.)

Legacy Systems/ An older technology or system related to a previous (usually outdated) computer system.

Liposomes/ A small vesicle made of cell membrane constituents that can be used as a drug delivery vehicle. (Maurya, SD. Liposome as a Drug Delivery Carrier- A Review. *Int Res J of Pharm* 2010; 1: 43-50.)

Lisp (Programming Language)/ Lisp, or LISP, is a high-level programming language initially specified in 1958 as part of the Dartmouth Summer Research Project on Artificial Intelligence by John McCarthy.

Logical Observation Identifiers Names and Codes (LOINC)/ A database and international standard for medical laboratory observations and measurements.

Logistic Regression/ A regression model where the dependent variable is categorical.

Long Short-Term Memory (LSTM)/ A simple recurrent neural network that can be built into a bigger recurrent neural network.

M/

Machine Learning/ A branch of artificial intelligence which allowing computers to learn from data via representation and generalization with uses in health care including analysis of the human genome and medical decision support. (Mani S et al. Medical Decision Support Using Machine Learning for Early Detection of Late-Onset Neonatal Sepsis. *J Am Med Inform Assoc* 2013; [Epub ahead of print]

Machine Vision (see Computer Vision)/ Technology and methodologies used to provide image-based analytics for inspection or control as well as robotic guidance in various industries.

Machine to Machine (M2M)/ Communication network that entails wireless devices forming an ecosystem that minimizes human interaction.

MapReduce (Apache)/ A programming paradigm that accommodates massive scalability of unstructured data across thousands of commodity clusters servers in an Apache Hadoop cluster.

Markov Model/ A system consists of a list of possible states of the system, possible paths, and rate parameters of transitions (also see hidden Markov model). (Sonnenberg FA et al. Markov Models in Medical Decision Making: A Practical Guide. *Med Decis Making* 1993; 13(4): 322-338.)

Mashup/ A Web application or derivative work that consists of combination of data of various sources. (Cheung KH et al. Semantic Mashup of Biomedical Data. *J of Biomed Inform* 2008; 41(5): 683-686.)

Massive Open Online Course (MOOC)/ On line distance education course and connection that is meant for a large scale participatory audience via Web access. (Harder B. Are MOOCs the Future of Medical Education? *BMJ* 2013; 346.)

MATLAB (Matrix Laboratory)/ A high level fourth generation language and interactive environment for numerical computation, visualization, and programming. (Singh SP et al. A Review of Estimating Development Time and Efforts of Software Projects by Using Neural Network and Fuzzy Logic in MATLAB. *Int J of Adv Res in Comp Sci and Sofware Eng* 2012; 2(10): 306-312.)

Meaningful Use/ In the use of EHR, meaningful use is for improving quality and care coordination as well as engaging patients maintaining privacy to lead to better clinical outcomes.

Mechanical Turk/ A mechanical chess playing automaton (but also an Amazon website to find and accept assignments using crowdsourcing).

Medical Decision Support (see Clinical Decision Support).

Medical Image Processing or Analysis/ Use of machine intelligence to perform quantitative analytics of medical imaging modalities such as CT, MRI, PET, or microscopy.

Medicare Access and CHIP Reauthorization Act (MACRA)/ A bipartisan legislation signed into law in 2015 that changes the way Medicare rewards clinicians for value over volume.

Medicine 2.0 (see Health 2.0)

Metadata/ Defined as data about data or how data is collected or formatted with relevance to data warehouses. (Sakai Y. Metadata for Evidence Based Medicine Resources. *Int Conf on Dublin Core and Metadata Applications*, 2001.)

Metabolomics/ The comprehensive characterization of small molecule metabolites in biological systems. (Madsen R et al. Chemometrics in Metabolomics- A Review in Human Disease Diagnosis. *Analytica Chimica Acta* 2010; 659(1-2): 23-33.)

mHealth (or m-health)(also Mobile Health)/ Practice of medicine and public health supported by mobile devices such as mobile phones or tablet computers. (Free C et al. The Effectiveness of M-Health Technologies for Improving Health and Health Services: A Systematic Review Protocol. *BMC Research Notes* 2010; 3: 250-258.)

Microbiome/ A discipline to examine how bacteria interact with each other and the human body to cause or prevent disease (Relman DA. Microbiology: Learning About Who We Are. *Nature* 2012; 486: 194-195.)

Microelectromechanical Systems (MEMS)/ Technology of microscopic devices especialy ones with moving parts.

Microfluidics (see Lab-on-a-Chip)

MicroRNA (miRNA)/ A short regulatory form of RNA that binds to a target RNA molecule and generally suppresses its translation by ribosomes. (Feero WG et al. Review Article: New Therapeutic Approaches to Mendelian Disorders. *N Engl J Med* 2010; 363: 852-863.)

Modified National Institute of Standards and Technology (MNIST)/ A large database of handwritten digits that is utilized as an image processing system training dataset.

Model-based Reasoning/ An artificial intelligence methodology in which an inference method is used based on a model of the physical world. (Croskerry P. A Universal Model of Diagnostic Reasoning. *Academic Med* 2009; 84(8): 1022-1028.)

Modifier Genes/ Genes that have a relatively small quantitative effect on the expression of another gene. (Nadeau JH. Modifier Genes in Mice and Humans. *Nature Reviews Genetics* 2001; 2: 16 5-174.)

MongoDB/ A free and open source platform that is a document-oriented NoSQL database program.

Monte Carlo Tree Search (MCTS)/ A heuristic search algorithm for certain types of decision processes such as in game playing for perfect information games (such as the game Go).

Moore's Law/ Observation by Gordon Moore, founder of Intel, that the number of transistors per square inch on integrated circuits double every two years since the circuit was invented.

Moravec's Paradox/ Observation by Hans Moravec, an AI robotics expert, that machines can solve things that are difficult for humans (such as geometrical problems) but things that are easy for humans are often difficult of machines.

Multiparameter Intelligent Monitoring in Intensive Care (MIMIC)/ Database that contains physiologic signals and vital signs time series from monitors as well as comprehensive clinical data for analytics and research.

Multiprotocol Label Switching (MPLS)/ A methodology used for WAN connectivity in order to route traffic inside a telecommunicatoins network.

Munging/ When referring to data, the process of cleaning and shaping the data prior to data mining; similar to data wrangling.

MYCIN/ An early expert system AI program that originated at Stanford University for treating blood infections.

N/

Naive Bayes Classifier/ A technique that is based on Bayesian theorem and is well suited for data with high input dimensionality.

Nanobots (also Nanorobots)/ Devices ranging in size from 0.1 to 10 micrometers and made of nanoscale or molecular components with promising use in biomedical technology. (Abeer S. Future Medicine: Nanomedicine. *Future* 2012; 25(3): 187-202.)

Nanomedicine/ Use of nanotechnology and nanomaterials for clinical applications such as in vivo contrast agents, drug carriers, or diagnostic devices. (Kim BYS et al. Current Concepts: Nanomedicine. *N Eng J Med* 2010; 363:2434-2443)

Nanotechnology/ The intentional design, characterization, production, and applications of materials, structures, devices, and systems by controlling their size and shape in the nanoscale range (1 to 100 nm). (Kim BYS et al. Current Concepts: Nanomedicine. *N Eng J Med* 2010; 363:2434-2443)

Nanotechnology, Biotechnology, Information technology, and Cognitive science (NBIC)/ Emerging technologies that follow a curve of exponential change. Another term is Genetics, Robotics, Internet, and Nanotechnology (GRIN) Technologies.

Nanotubes (see Fullerene)

Natural Language Processing (NLP)/ A field in artificial intelligence and computer science to study the interaction between human natural language and computers. (Huske-Kraus D. Text Generation in Clinical Medicine- A Review. *Methods of Inform in Med* 2003; 42(1): 51-60.)

Neat and Scruffy AI/ The two philosophies of AI: Neats prefer solutions that are elegant and clear whereas scruffies believe that intelligence is very complex.

Neo4j (Neo4j)/ A graph database management system with native graph storage and processing that is presently the most popular graph database.

Net Neutrality/ The basic principle that disallows internet service providers from altering any content by speeding up/slowing down or by blocking any content or applications one like to use; the neutrality keeps internet free and open.

Network File System (NFS)/ A file system that enables storage and retrieval of data from files from multiple sources/

Neural Network or Nets (see Artificial Neural Network).

Neural Lace (Neuralink)/ Elon Musk's concept of a brain-computer system what would link human brains with a computer interface to achieve symbiosis.

Neuromorphic Chip/ Silicon chips that are modeled on biological brains designed to process sensory data as a human brain without specific programming.

Neuromorphic Computing/ Design of computers that will possess three characteristics that brains have (but computers do not): low power consumption, fault tolerance, lack of need to be programmed. (Neuromorphic Computing: The Machine of New Soul. *The Economist*, 2013).

Neuromorphic Engineering/ (see Neuromorphic Computing)

Next Generation Sequencing (NGS)/ The inexpensive production of large volumes of sequence data that holds an advantage over the first generation automated Sanger sequencing technique. (Metzker ML. Sequencing Technologies- The Next Generation. *Nature Reviews (Genetics)* 2010; 11: 31-42.)

NLP (see Natural Language Processing).

O/

One-Shot Learning/ A machine learning strategy promulgated by Stanford machine learning expert Li Fei Fei to learn information about object categories from one or very few training images instead of large amounts of such data.

Online Analytical Processing (OLAP)/ An approach used in business intelligence to answer analytical query that is multi-dimensional (relational database, reports, and data mining).

Online Transaction Processing (OLTP)/ A class of information systems that deal with small and interactive transactions that require very low response times (as opposed to batch processing).

Optical Character Recognition (OCR)/ Electronic conversion and information entry as well as recognition of images of text (typed or handwritten) to a machine-encoded text.

Organ Printing/ Biomedical application of rapid prototyping, or additive layer-by-layer biomanufacturing, is an emerging *in situ* transforming technology that is new compared to traditional solid scaffold-based tissue engineering. (Mironov V et al. *Regenerative Medicine* 2008; 3(1): 93-103.)

Overfitting/ A type of modeling error in which the function or learning system is too closely fitting a training data set as to not able to accurately predict outcomes of the untrained data set (describes the noise instead of the underlying concept).

P/

P4 (Predictive, Personalized, Preventive, and Participatory) Medicine/ Philosophy to allow biotechnology to manage an individual's health and wellness instead of disease in a personalized approach. (Hood L et al. Predictive, Personalized, Preventive, Participatory (P4) Cancer Medicine. *Nature Reviews Clinical Oncology* 2011; 8: 184-187).

Parsing/ An analysis used in NLP of a string of symbols in natural or computer language following the rules of grammar.

Part-of-Speech Tagging (POST)/ The natural language process in which the words in a sentence are marked to correspond to part of speech.

Patient-Centered Outcome Research Institute (PCORI)/ Organization that helps to make informed healthcare decisions and better outcomes by promoting evidence-based information.

Perceptron/ A machine learning algorithm for supervised learning of binary classifiers initially conceived in 1957 by Frank Rosenblatt and is the basis for artificial neural network.

Personalized Medicine/ A more precise and customized extension of traditional approaches to understanding and treating disease that has advanced due to wide availability of genetic information. (Burke W et al. Personalized Medicine in the Era of Genomics. *JAMA* 2007; 298(14): 1682-1684.)

Petabyte/ One quadrillion or 10^{15} bytes which is enough storage capacity to store the DNA of all Americans (1000 Terabytes = 1 Petabyte and 1000 Petabyte = 1 Exabyte).

Pharmacogenetics/ The study of how the actions of and reactions to drugs are dependent on the variations of an individual's genes and metabolic pathways. (Hamburg MA et al. The Path to Personalized Medicine. *N Engl J Med* 2010; 363(4): 301-304.)

Pharmacogenomics/ A term at times confused with Pharmacogenetics, pharmacogenomics is the genomic discipline to study genes in the entire genome across the population that influence drug response (the premise for personalized medicine). (Brownstein C et al. Integration of a Standardized Pharmacogenomic Platform for Clinical Decision Support at Boston Children's Hospital. *BMC Proceedings* 2012; 6:P5.)

Phenomics/ Use of large-scale, high-throughput assays and bioinformatics approaches to investigate how genetic instructions actually translate into tangible phenotypic traits. (Kodituwakku PW et al. From Research to Practice: An Integrative Framework for the Development of Interventions for Children with Fetal Alcohol Spectrum Disorders. *Neuropsychol Rev* 2011, 21: 204-233.)

Picture Archiving and Communication System (PACS)/ Medical imaging technology for storing as well as presenting medical images by a myriad of imaging modalities (MRI, X-Ray, CT, and Ultrasound).

Pixel/ It is considered the smallest element or a sample of a picture in digital imaging.

Pluripotency (also Induced Pluripotent Stem Cells)/ The potential of a cell to develop into more than one type of mature cell, usually any of the three germ layers (endoderm, mesoderm, or ectoderm). (Wu SM et al. Harnessing the Potential of Induced Pluripotent Stem Cells for Regenerative Medicine. *Nature Cell Biology* 2011; 13: 497-505.)

Podcast/ A multimedia digital file that is accessible on the Internet and downloadable to a computer or other device with the potential for medical education and patient education. (Savel RH et al. The iCritical Care Podcast: A Novel Medium for Critical Care Communication and Education. *J Am Med Inform Assoc* 2007; 14: 94-99.)

Polanyi's Paradox/ The observation that our intuitive tacit knowledge of how the world works (bottom of iceberg) often exceeds our explicit understanding (tip of the iceberg) was forwarded by the polymath and philosopher Michael Polanyi.

Posterior Probability/ In Bayes' rule, it is a conditional probability of a given event that is computed after observing a second event.

Precision Medicine/ The coupling of clinical-pathological profiles with molecular profiles to tailor the diagnostic and therapeutic regimen to the individual. (Mirnezami R et al. Preparing for Precision Medicine. N Engl J med 2012; 366: 489-491.)

Predictive Analytics/ The capability to predict future human behavior or events by data mining and machine learning techniques (see above). (Siegel E. *Predictive Analytics: The Power to Predict Who Will Click, Buy, Lie, or Die.* John Wiley and Sons, Hoboken, 2013).

Predictive Modeling/ This strategy uses statistics to predict outcomes via one or more classifiers.

Preimplantation Genetic Diagnosis (PGD)(also Preimplantation Genetic Screening, PGS)/ The genetic profiling of the embryo prior to implantation and as part of *in vitro* fertilization. (Harper JC et al. Preimplantation Genetic Diagnosis: State of the Art 2011. *Hum Genet* 2011; 131(2): 175-186.)

Principal Components Analysis (PCA)/ A simple but powerful statistical technique that is useful in extracting relevant information from confusing data sets for such applications as face recognition and image compression.

Probabilistic Reasoning/ Usage of past situations and examples and apply statistics to predict a likely outcome.

Protected Health Information (PHI)/ Information regarding health status and care as well as payment that can be associated with an individual.

Proteomics/ Study of structure and function of proteins. (Yeager ME et al. Plasma Proteomics of Differential Outcome to Long-Term Therapy in Children with Idiopathic Pulmonary Arterial Hypertension. *Proteomics Clin Appl* 2012; 6(5-6): 257-267.)

Python/ A powerful (multi-paradigm) dynamic programming language that is used in a myriad of application domains with certain advantages for research: language interoperability, data structures, available libraries, and balance of high and low level programming.

Pytorch (Python)/ A Python-based library used as a deep learning developmental framework for a efficient experimentation.

Q/

Quantifiable Self/ Self-knowledge of medical conditions or disease through self-tracking. (Swan M. Health 2050: The Realization of Personalized Medicine Through Crowdsourcing, the Quantified Self, and the Participatory Biocitizen. *J Pers Med* 2012; 2: 93-118.)

Quantum Bit (or Qubit)/ The basic unit of information in quantum computing that is analogous to the bit in classical computing.

Quantum Computing/ Computing that will harness the power of atoms and molecules to perform memory and processing tasks with quantum bits, or qubits.

Quantum Dot/ Materials that consist of a core-and-shell structure (e.g. CdSe coated with zinc and sulfide with a stabilizing molecule and a polymer layer coated with a protein. (Kim BYS et al. Current Concepts: Nanomedicine. *N Eng J Med* 2010; 363:2434-2443.)

Qubits (see Qubit).

Quick Response (QR) Code/ A matrix barcode that is attached to an item which has information related to the item. (Samwald M et al. Pharmacogenomics in the Pocket of Every Patient? A Prototype Based on Quick Response Codes. *J Am Med Inform Assoc* 2013; 20: 409-412.

R/

R/ A computer programming language and environment for statistical computing and high quality graphics particularly well suited for biomedical sciences.

Radio-frequency Identification (RFID)/ Usage of radio waves to communicate between the reader and an electronic tag that is affixed to an object or person. (Kolokathi A et al. Radio Frequency Identificatoin (RFID) in Healthcare: A Literature Review. *Stud Health Technol Inform* 2013: 190: 157-159.)

Random Forest/ An ensemble of decision trees (with each tree constructed by using a random subset of training data) that will output a prediction value.

Rectified Linear Unit (ReLU)(or Rectifier)/ ReLU is the most used activation function in deep learning. An activation function is a node added to the output end of any neural netowrk.

Recurrent Neural Network (RNN) or Recurrent Nets/ A type of deep learning that is useful in sequential data such as text and speech as well as time series data.

Recursive Cortical Network (RCN)/ A generative object-based model that deviates from deep learning in that it begins with a scaffold (rather than tabla rasa) and therefore is much more data efficient.

Regenerative Medicine/ Field of biomedical science that utilizes technology to use stem cells to rejuvenate, replace, and regenerate body cells, tissue and organs. (Gurtner GE et al. Progress and Potential for Regenerative Medicine. *Ann Rev of Med* 2007; 58: 299-312.)

Reinforcement Learning/ Machine learning that is inspired by behavioral psychology to maximize cumulative reward.

Relational Database/ A database with a relational model of data and the software systems that maintain these databases are called relational database management system (RDBMS).

ResearchKit (Apple)/ An open source framework introduced by Apple that provides an opportunity for researchers and developers to create apps for medical research.

Resource Description Framework (RDF)/ A family of World Wide Web Consortium specifications for metadata data model and are documents that are used ro store information in the RDF language.

Rich Site Summary (RSS)(also Really Simple Syndication)/ A format for delivery of regularly changing web content and a family of web feed formats to gather updated information including blog entries, audio and video information, and news articles and to present a summarized text for an update. (McLean R et al. The Effect of Web 2.0 on the Future of Medical Practice and Education: Darwikinian Evolution or Folksonomic Revolution? *Med J Aust* 2007; 187(3): 174-177.)

Robotic Process Automation (RPA)/ Intelligent software robots deployed to automate repetitive activities with the user interface of a computer system.

Robots/ Technology that entails design, construction, maintenance, and application of robots with their computer environment; health care applications in children include rehabilitation, exoskeletons, and virtual visits. (Blazquez MP. Clinical Applications of Robotics in Children with Cerebral Palsy. *Biosystems and Biorobotics* 2013; 1: 1097-1102.)

Robotic Surgery/ Robotic systems (the Intuitive Surgical da Vinci Surgical System and the Zeus MicroWrist Surgical System) used to assist in the performance of certain surgical procedures, especially laparoscopic procedures. (van Haasteren G et al. Pediatric Robotic Surgery: Early Assessment. *Pediatrics* 2009; 124(6): 1642-1649.)

Rule-based Reasoning/ An artificial intelligence technique (under the knowledge-based systems) that involves rules, data base, and interpreter for the rules. (Kumar K et al. Hybrid Approach Using Case-Based Reasoning and Rule-Based Reasoning for Domain Independent Clinical Decision Support in ICU. *Expert Systems with Applications* 2009; 36(1): 65-71.

S/

Scruffy (see Neat and Scruffy).

Segmentation/ The process of dividing or partitioning a medical image in computer vision into many segments (called pixels) for analysis.

Self-Organizing Map/ A data visualization technique to display an ANN that uses unsupervised learning.

Semantic (or Smart) Data Lake/ Use of semantic graph models as the format for storing data for more efficient use of data during retrieval (Data Reservoir?- ACC)

Semantic Net or Network/ A knowledge representation methodology used for propositional information (or mathematically a labelled directed graph).

Semantic Web/ Extension of the World Wide Web with metadata that will allow users to share content beyond the traditional boundaries of applications and websites and for computers to "talk" to each other; also termed "Web 3.0". (Giustini D. Web 3.0 and Medicine. *BMJ* 2007; 335: 1273-1275.)

Sentence Compression Algorithm/ Strategy to extract the best answer for a query to be displayed in the featured snippets.

Sentient AI/ AI that possesses properties of self-awareness.

Signal Processing/ Technology that uses mathematical and computational representation for transferring information contained in various formats.

Single (Simple) Nucleotide Polymorphisms (SNPs)/ A single-nucleotide variation in a genetic sequence that is a common form of variation in the human genome. (Manolio TA. Genomewide Association Studies and Assessment of the Risk of Disease. *N Engl J Med* 2010; 363: 166-176.)

Singularity/ This refers to the point in time (estimated to be around the year 2045) during which technological intelligence supercedes human intelligence. This is a concept initially attributed to mathematician John von Neumann but popularized by the science fiction writer Vernor Vinge and the futurist Ray Kurzweil. Also known as technological singularity. (Kurzweil, Ray. *The Singularity is Near*. The Penquin Group, New York, 2005.)

Small (or Short) Interfering RNA (siRNA)/ A short, double-stranded regulatory RNA molecule that binds to and induces the degradation of target RNA molecules. (Hudson KL. Review Article: Genomics, Health Care, and Society. *N Engl J Med* 2011; 365: 1033-1041.)

Small Iron Oxide Nanoparticles (SPIONS)/ These nanoparticles contain one domain that leads to a relatively large generated magnetic field (Kim BYS et al. Current Concepts: Nanomedicine. *N Eng J Med* 2010; 363:2434-2443.)

Smart Wearable Systems (SWS)/ Devices ranging from sensors and actuators to other monitoring devices for management of patients' health status. (Chan M et al. Smart Wearable Systems: Current Status and Future Challenges. *Artif Intell Med* 2012; 56(3): 137-156.)

SNOMED CT (see Systemized Nomenclature of Medicine- Clinical Terms).

Social Media/ The creation, sharing, and exchange of information and ideas in the virtual community; includes tools such as Twitter, Facebook, YouTube, Foursquare, and other social tools. (Jain SH. Practicing Medicine in the Age of Facebook. *N Engl J Med* 2009; 361: 649-651.)

Software as a Service (SaaS)/ A distribution model in which software is made available to the client via a provider hosts applications over the internet.

Software Agent/ A computer program capable of performing without direct supervision and is the computer analog of an autonomous robot.

Spark (Apache)/ An open source big data processing engine with the ecosystem consisting of structured data, streaming analytics, machine learning, and graph computation.

SPARQL Protocol and RDF Query Language (SPARQL)/ A semantic query language for databases stored in RDF format.

Spiking Neural Network (SNN)/ A third-generation neural network that bridges the gap between machine learning and neuroscience by using neuronal properties. SNN uses discrete events rather than continuous values that conventional machine or deep learning use.

SQL/ A domain-specific special purpose language used in programming and designed for managing data in a relational database management system (RDMS).

Stem Cell (see Pluripotency and Embryonic Stem Cell)/ Unspecialized cells capable of renewing via cell division and also capable of being induced to become tissue specific cells. (Uccelli A et al. Mesenchymal Stem Cells in Health and Disease. *Nature Reviews Immunology* 2008; 8: 726-736.)

String/ A sequence of characters in computer programming.

Strong AI (or Artificial General Intelligence, AGI)/ An AI machine that is at least or more skillful or flexible as a human with machine intelligence that can think and function similar to humans.

Structured Query Language (SQL)/ A special-purpose query programming language for accessing and managing databases organized in a relational database management system (RDBMS). (Das AK et al. A Temporal Query System for Protocol-Directed Decision Support. *Methods of Information in Medicine* 1994; 33(4): 358-370.)

Superintelligence/ A form of intelligence that supercedes the intelligence of humans and is associated with technological singularity.

Supervised Learning/ Machine learning of predicting from labeled training data with each example being an input object and a desired output.

Support Vector Machines (SVM)/ A machine learning technique in which supervised learning models with learning algorithms are used that analyze data for classification and regression analysis.

Swarm Intelligence/ Collective intelligence of a decentralized but self-organizing system that was introduced in the context of cellular robotic systems.

Syllogism/ A form of reasoning (deductive) in which a conclusion is drawn from two given or assumed propositions.

Systemized Nomenclature of Medicine- Clinical Terms (SNOMED CT)/ A suite of designated standards of clinical terminology used by clinicians and others in electronic health and medical records.

Systems Biology/ Integration of biology and medicine along with technology and computation as a discipline to study biological components from molecules to organisms or entire species. (Minoo P et al. Systems Biology and Pediatric Research. *Ped Res* 2013; 73: 499-501.)

T/

Tag Cloud (see Word Cloud)

Telehealth/ Health care services (with a wider spectrum than telemedicine) being delivered via telecommunication with promise to increase access to specialized services (Farmer JE. Telehealth for Children with Special Health Care Needs: Promoting Comprehensive Systems of Care. *Clin Pediatr* 2001; 40(2): 93-98.)

Telemedicine/ The use of medical information exchanged from one site to another site via electronic communications. (Karp WB et al. Use of Telemedicine for Children with Special Care Needs. *Pediatrics* 2000; 105(4): 843-847.)

Tensor Processing Unit (TPU)(Google)/ Machine learning chips, more powerful than CPUs or GPUs, designed specifically to increase the speed of machine learning tasks.

TensorFlow™ (Google)/ An open source software library for machine intelligence developed by Google Brain team for numerical computation using data flow graphs as well as deep learning.

Terabyte (TB)/ Computer memory storage capacity that is a trillion bytes or a thousand gigabytes; 10 terabytes can hold the entire literary collection of the Library of Congress (1000 Terabytes = 1 Petabyte and 1000 Petabyte = 1 Exabyte).

Text Mining (Text Analytics or Text Data Mining)/ The process of extracting high-quality information from natural language text.

Tissue Engineering (see Regenerative Medicine)/ The use of cells and materials to improve or replace biological tissue or organ. (Bianco P et al. Stem Cells in Tissue Engineering. *Nature* 2001; 414: 118-121.)

Tokenization/ The process of breaking up a stream of text into words or phrases called tokens which can be used to secure information for encryption.

Transfer Learning/ Machine learning type of learning that focuses on storing knowledge gained from solving one problem and then applying this knowledge to solve another problem.

Transhumanism/ Movement with philosophy of improving humans with available technological modifications from intellectual and physical perspectives. (Lucas MS. Baby Steps to Superintelligence: Neuroprosthetics and Children. *J Evol Technol* 2012; 22(1): 132-145.)

Transmission Control Protocol/ Internet Protocol (TCP/IP)/ A combination of two sets of rules or protocols (HTTP and HTTPS for non-secure and secure data transmissions are examples) governing communications (in the form of packets) amongst the computers on the Internet.

Twitter/ An online social networking and microblogging service that utilizes “tweets” or 140-character text messages as its main mode of communication. (Chretien KC et al. Physicians on Twitter. *JAMA* 2011; 305(6): 566-568.)

Turing Test/ A test of artificial intelligence devised by Alan Turing, the famous mathematician, wherein a human interrogator is given the task of distinguishing between a human and a computer based on the replies to questions. While virtual assistants and IBM Watson would not be able to pass this test, the AI program Eugene Goostman did pss the Turing test in June of 2014. (Chandrasekaran B. On Evaluating AI Systems for Medical Diagnosis. *AI Magazine* 1983; 4(2): 34-48.)

U/

Unsupervised Learning/ Type of machine learning in which a function is predicted based on unlabeled data so there is no error/reward signal for the predicted solution.

Unmanned Aerial Vehicle (UAV)(see drone).

Underfitting/ Phenomenon when a statistical model or machine learning algorithm fails to fit the data sufficiently so that the model is excessively simple.

User Experience (UX)/ The optimization of the usability of the product or service from the user's perspective (human-first) and is more analytical and technical than UI (something that looks great but not feel good to use is good UI but not good UX).

User Interface (UI)/ The graphic design of the product or service and at times confused with UX (something that feels good to use but not good to look at is good UX but not good UI).

.

V/

Variants of Uncertain Significance (VUS) (also Allelic Variant of Unknown Significance)/ An alteration in the normal sequence of a gene that has unknown clinical significance and disease risk. (McCarthy MI et al. Genome-wide Association Studies for Complex Traits: Consensus, Uncertainty, and Challenges. *Nature Reviews Genetics* 2008; 9: 356-369.

Visual Analytics/ The science of analytical reasoning supported by interactive visual interfaces and consists of an integral approach combining visualization, human factors, and data analysis. (Caban JJ et al. *2011 Workshop on Visual Analytics in Healthcare: Understanding the Physician Perspective* 2012; 2(1): 29-31.)

Virtual AI Assistant (VA)/ A sentient digital assistant that uses artificial intelligence for facilitating one's digital life or other activities and tasks (e.g. Apple's Siri, Microsoft's Cortana, Amazon's Echo, Facebook's M, or Google's Now).

Virtual Private Networks (VPN)/ A private network extension to connect a shared or public network to enable organizations and individuals to transmit data between computers.

Virtual Reality (VR)(see Augmented Reality)/ Technology that uses images and sounds in order to simulate a user's physical presence in a virtual environment.

Voxel/ A value in three-dimensional space (as opposed to pixel in two-dimensional picture).

W/

Watson (IBM)/ A supercomputer that utilizes a portfolio of natural language processing, information retrieval, knowledge representation, and machine learning with four terabytes of disk storage to be able to read close to 100 million pages per second and defeat human contestants on the quiz show *Jeopardy!*. (Yuan MJ. Watson and Healthcare: How Natural Language and Semantic Search Could Revolutionize Clinical Decision Support. IBM 2011)

Wearable Technology or Devices/ Devices that can monitor vital signs (such as heart rate, blood pressure, or pulse oximetry) and electrocardiogram (for waveform analysis and heart rhythm assessment).

Weak AI (also Narrow AI)/ AI or machine intelligence that is based on a single focused task (such as chess playing or Go playing).

Web 2.0/ A term describing new collaborative Internet applications with key elements include: RSS (see above), blogs, wikis, and podcasts. (McLean R et al. The Effect of Web 2.0 on the Future of Medical Practice and Education: Darwikinian Evolution or Folksonomic Revolution? *Med J Aust* 2007; 187(3): 174-177.)

Web 3.0 (see Semantic Web)/ A term to describe the evolution of the web in finding and organizing new information beyond the boundaries of websites. (Giustini D. Web 3.0 and Medicine. *BMJ* 2007; 335: 1273-1275.)

Wide Area Network (WAN)/ A computer or telecommunications network over a large geographical area.

Wiki (or Medical Wiki)/ A website or online resource that allows certain users to add and edit medical information as a collective group. (Bender JL et al. Collaborative Authoring: A Case Study of the Use of a Wiki As a Tool to Keep Systematic Reviews Up to Date. *Open Med* 2011; 5(4): e201-e208.)

Wireless Sensor Networks (WSN)/ A network of autonomous sensors for monitoring conditions in an environment but has potential healthcare applications. (Alemdar H et al. Wireless Sensor Networks for Healthcare: A Survey. *Computer Networks* 2010; 54(15): 2688-2710.)

Wolfram Alpha (or WolframAlpha) / A powerful computational knowledge engine that uses natural language processing and answers questions (similar to DeepQA project by IBM). (Shah NP. Recent Technological Advances in Natural Language Processing and Artificial Intelligence. *arXiv preprint arXiv* 2012; 1208:4079.)

Word Cloud (also Tag Cloud)/ A method of representing text data to visualize keyword metadata on websites (by size of font and by color). (McGee RG et al. A Picture is Worth a Thousand Words. *Am J of Transpl* 2011; 11(4): 871-872.)

X/

XAI (or Explainable AI)/ AI, particularly machine and deep learning, that can be understood and trusted by human participants in applications of AI.

XML/ see Extensible Markup Language.

Y/

YouTube/ An online video-sharing website with video clips uploaded by individuals or by media corporations and even hospitals. (Stamelou M et al. Movement Disorders on YouTube- Caveat Spectator. *N Engl J Med* 2011; 365: 1160-1161.)

Z/

Zettabytes (ZB)/ One sextillion or 10^{21} bytes (or one billion terabytes) with the total amount of global data around 3 zettabytes.