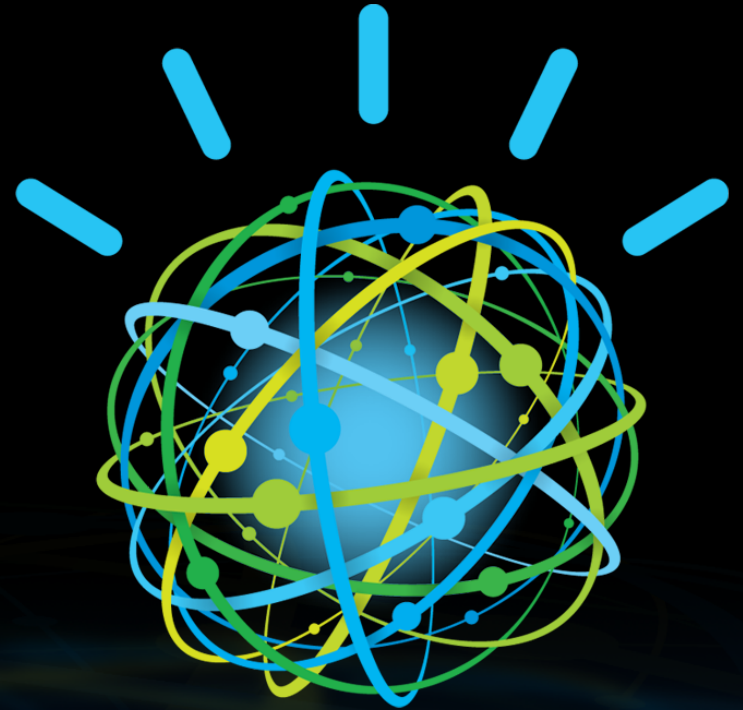# Big Data and IBM Watson

Making a Market, Making a Difference

Topics

- What is Big Data and why does it matter?

- How does IBM Watson leverage Big Data Analytics in Healthcare?

- What are some of the early results from Watson customers?

Four years ago, we started working with organizations to build a smarter planet.

Through thousands of client engagements, we learned that analytics is fundamental to success.

Data is the rapidly becoming the foundation for a Smarter Planet

Cloud

Mobile

Social

Big Data

Internet of Things

Outperformers are capitalizing on the opportunity

**% of CEO's** who believe their organization is good at driving value from data

Underperformers

Outperformers

Access to data

26%

54%

110%

Draw insights from data

26%

54%

108%

Translate insight into action

31%

57%

86% more

Source: IBM CEO Study, 2012. Q22 "How good is your organization at driving value from data? [Today]" (n=631 to 636)

CIOs are turning to innovative technologies to deliver outcomes

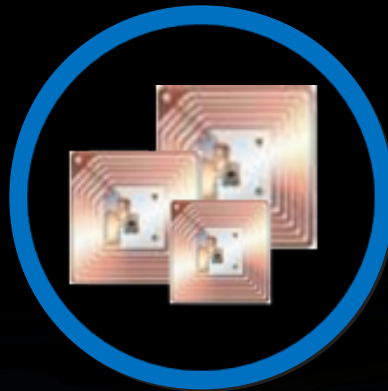## Technologies for innovation and solving specific business challenges

Technology	Percentage
Big Data Analytics	83%
Mobile	74%
Cloud	68%
BPM	60%
Security Intelligence	58%
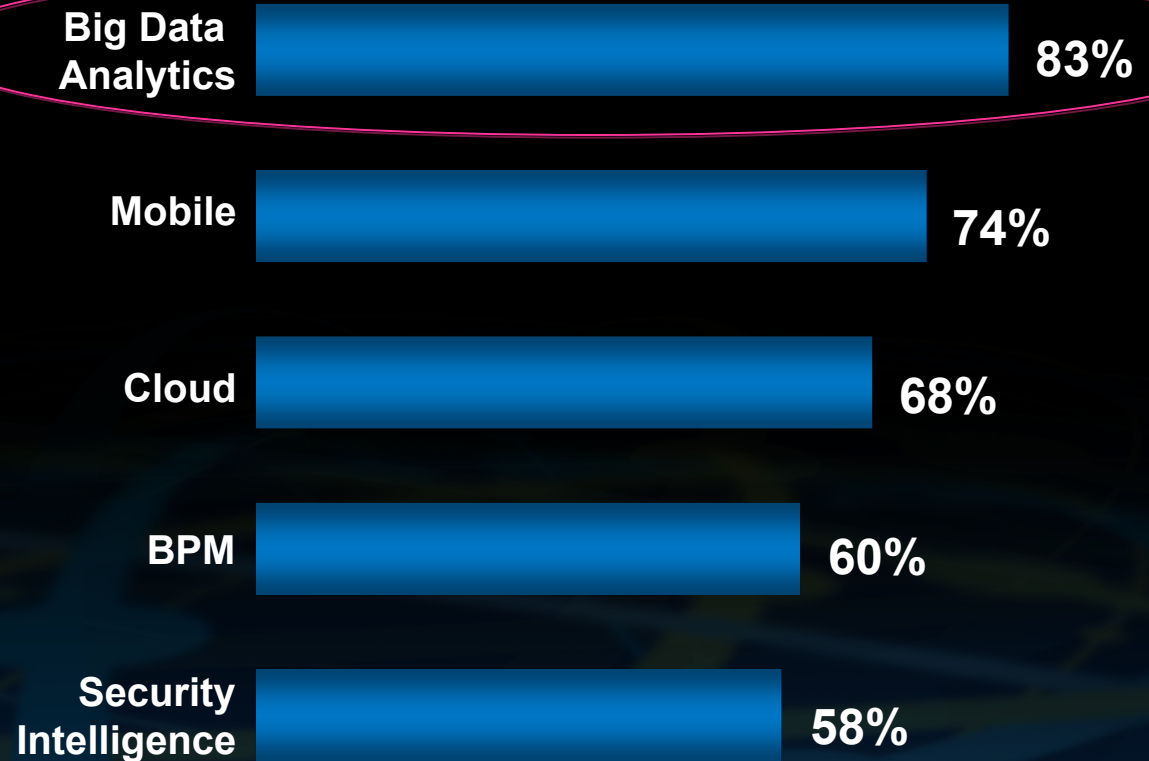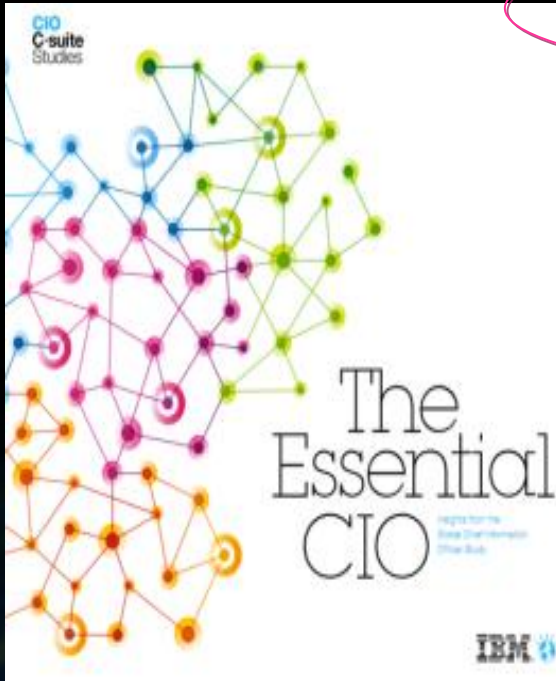
Big Data: Think beyond the traditional data types

Transactional & Application Data	Machine Data	Social Data	Enterprise Content

Transactional & Application Data	Machine Data	Social Data	Enterprise Content
• Volume	• Velocity	• Variety	• Variety
• Structured	• Semi-structured	• Highly unstructured	• Highly unstructured
• Throughput	• Ingestion	• Veracity	• Volume

Big Data is characterized by the 'Four Vs'
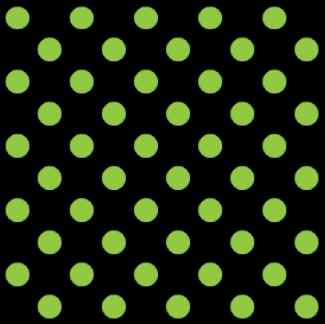
Volume	Velocity	Variety	Veracity*
Data at Rest	**Data in Motion**	**Data in Many Forms**	**Data in Doubt**
Terabytes to exabytes of existing data to process	Streaming data, milliseconds to seconds to respond	Structured, unstructured, text, multimedia	Data inconsistency & incompleteness, ambiguities, latency, deception, model approximations
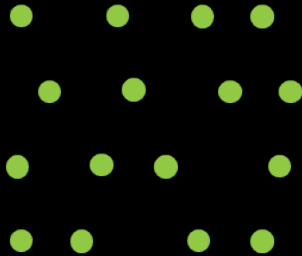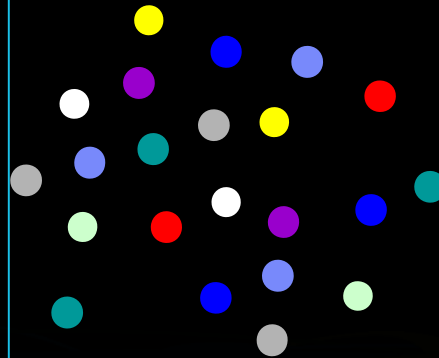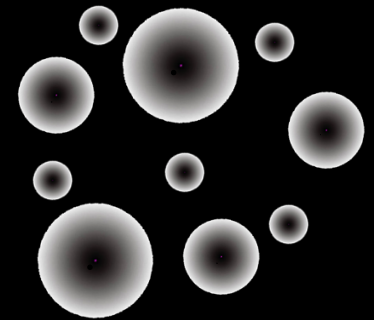
* Truthfulness, accuracy or precision, correctness

Big data creates more opportunities for analytics

Gain more complete answers
Extend existing analytics to provide additional insights

Using information and analytics in new ways and exciting ways

Create new perspectives
Extend analytics to communities and processes not reached before

Reduce IT Costs
Rethink existing approaches to how data and content is managed, stored and analyzed to reduce infrastructure costs

Uncover new business opportunities
Identify new offerings and new business models that create value

Web & Social Interaction Data

Multimedia Data

Text, Content, & Documents

Storage & Network Data

Transactional Data

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

Medical information is doubling every 5 years, much of which is unstructured

81% of physicians report spending 5 hours or less per month reading medical journals

1 in 5
diagnosis that are estimated to be inaccurate or incomplete

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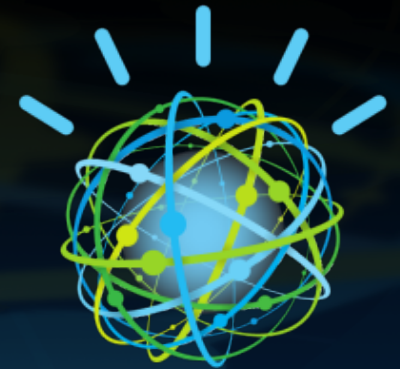errors in the way medications are prescribed, delivered and taken in the U.S. every year

44,000 -98,000
# of Americans who die each year from preventable medical errors in hospitals alone

Topics

- What is Big Data and why does Big Data Analytics matter?

- How does IBM Watson leverage Big Data Analytics in Healthcare?

- What are some of the early results from Watson customers?

Brief History of IBM Watson

IBM
Research
Project
(2006 –)

R&D

Jeopardy!
Grand
Challenge
(Feb 2011)

Demonstration

Watson
for
Healthcare
(Aug 2011 –)

Commercialization

New IBM Division

Watson
for Financial
Services
(Mar 2012 –)

Expansion

Watson
Industry
Solutions
(2012 –)

Cross-industry
Applications

Why Watson for healthcare?

Complexity
- Diagnosis and treatment errors
- Shortage of MDs
- Demand for remote medicine

Policy Changes
- Shift from Fee-for-Service to ACOs
- Focus on Wellness and Prevention
- Universal coverage

Evidence-based Medicine

Personalized Medicine

Costs
- Costs are 18% of US GDP
- 34% of $2.3T US spend is waste
- Costs can vary up to 10x

Info Overload
- Medical data doubles every 5 years
- Detailed patient biomedical markers
- Targeted therapies

Why Cancer Care Solutions?
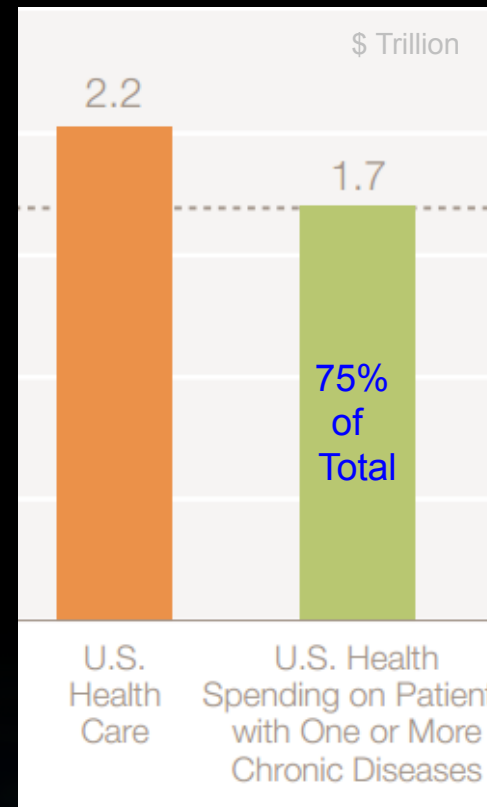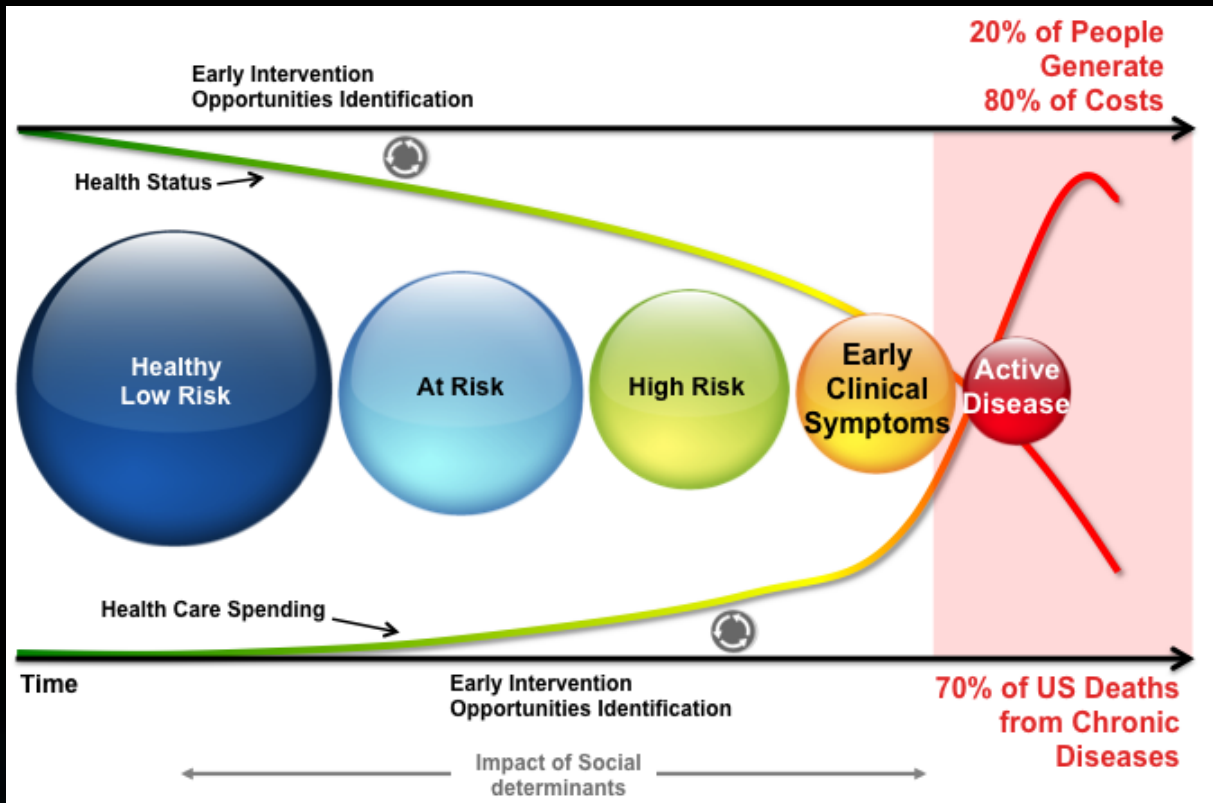
$263.8B
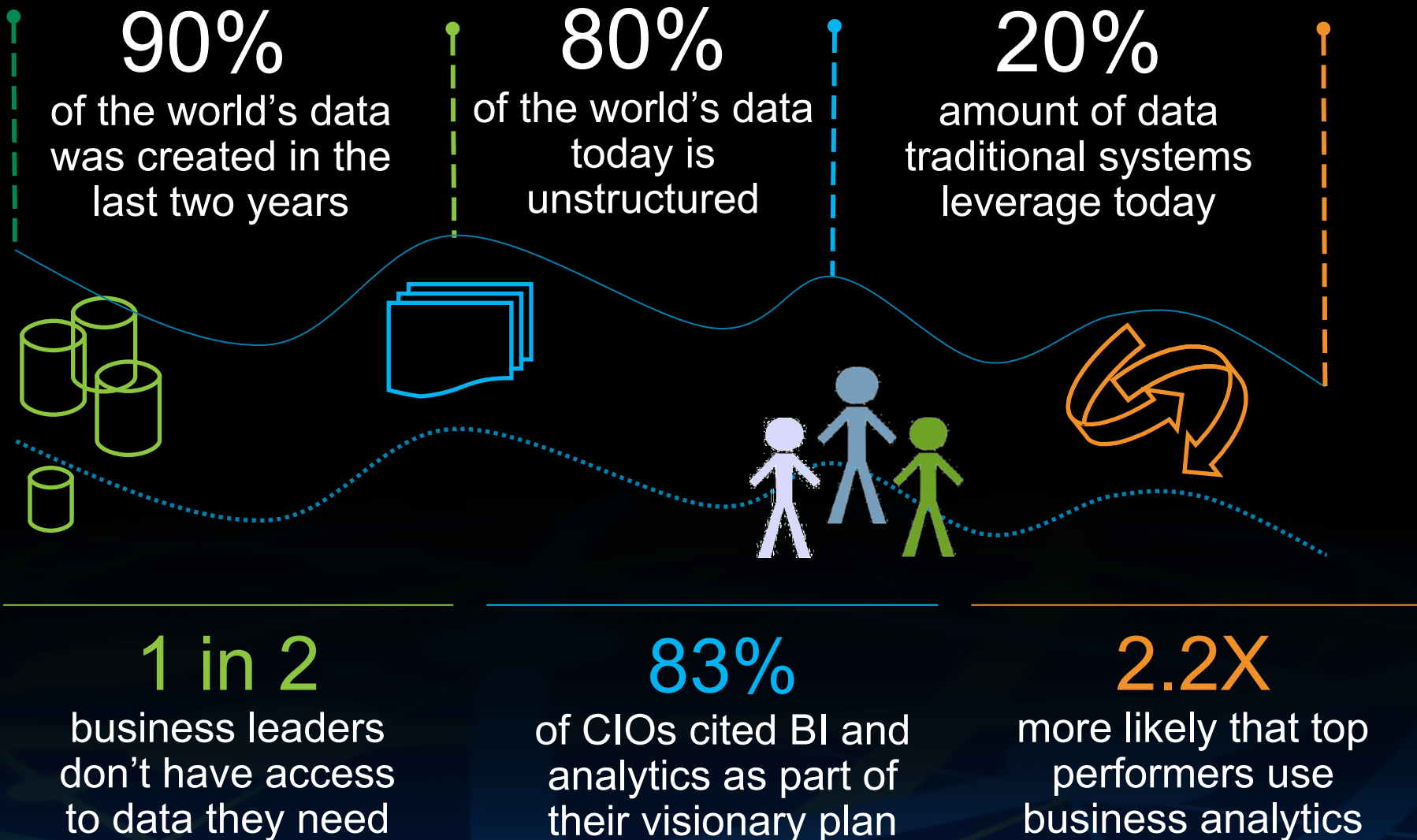
overall costs of
cancer
in the US in 2010

3X

rate cancer cost climbs
vs. std. health costs or
15-18% / yr.

20-44%%

of cancer cases
receive the wrong
diagnosis initially

Businesses are "dying of thirst in an ocean of data"

90% of the world's data was created in the last two years

80% of the world's data today is unstructured

20% amount of data traditional systems leverage today

1 in 2 business leaders don't have access to data they need

83% of CIOs cited BI and analytics as part of their visionary plan

2.2X more likely that top performers use business analytics

Why is it so hard for computers to understand us?

Welch ran this?

Person	Organization
L. Gerstner	IBM
J. Welch	GE
W. Gates	Microsoft

"If leadership is an art then surely Jack Welch has proved himself a master painter during his tenure at GE."

- Noses that run and feet that smell?
- How can a house burn up as it burns down?
- Does CPD represent a complex comorbidity of lung cancer?
- What mix of zero-coupon, non-callable, A+ munis fit my risk tolerance?

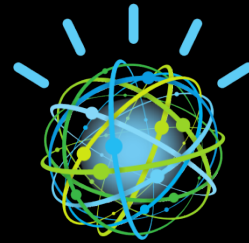# IBM Watson combines transformational technologies

2 Generates and evaluates evidence-based hypothesis

1 Understands natural language and human communication
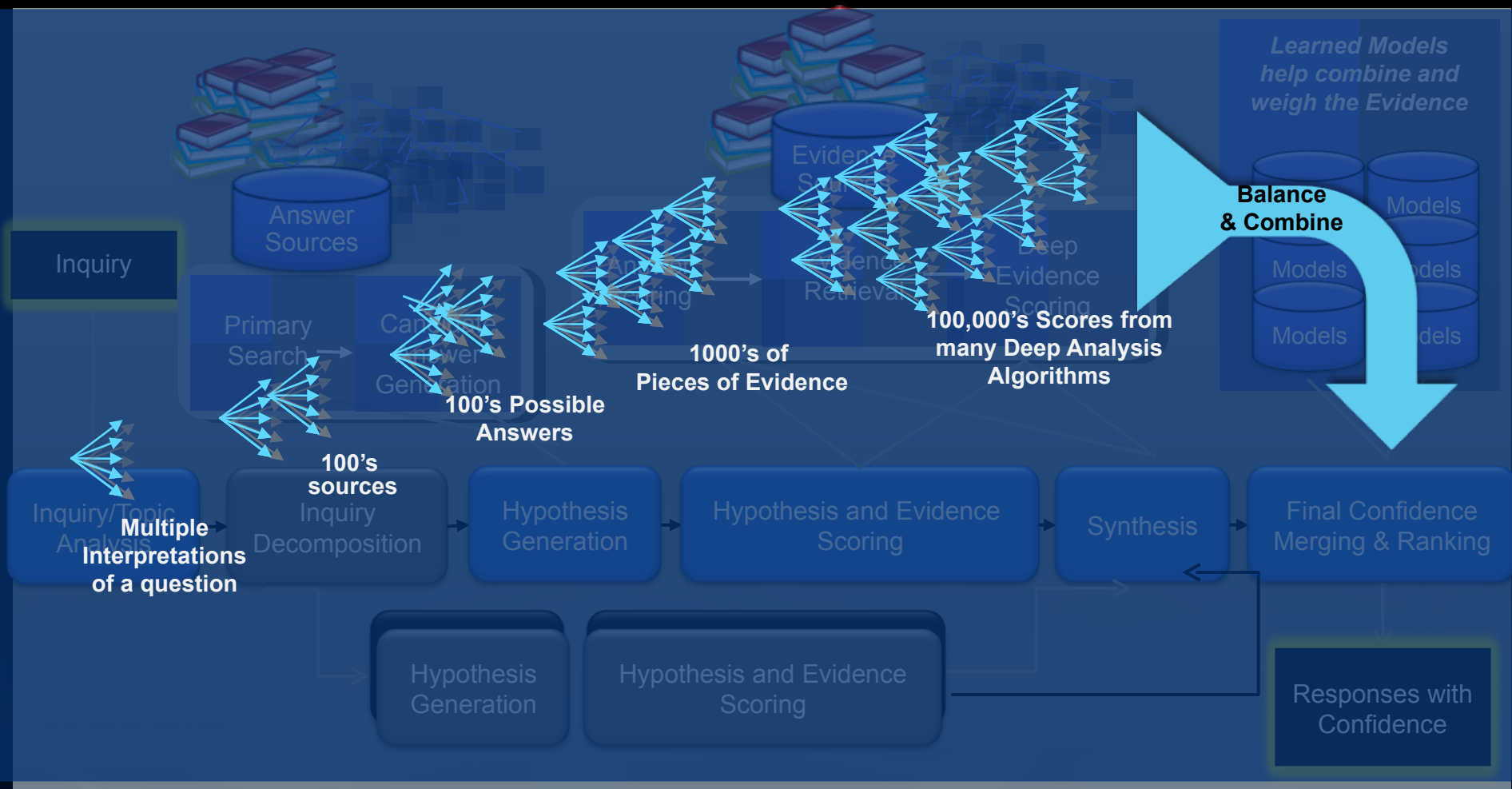
3 Adapts and learns from user selections and responses

...built on a massively parallel architecture optimized for IBM POWER7

How Watson Works: DeepQA Architecture

# Putting the pieces together at point of impact can be game changing

Diagnosis Models

Symptoms · Fam. History · Pat. History · Medications · Findings

Confidence

Symptoms
- difficulty swallowing
- fever
- dry mouth
- thirst
- anorexia
- frequent urination
- dizziness
- no abdominal pain
- no back pain
- no cough
- no diarrhea

Family History
- Oral cancer
- Bladder cancer
- Hemochromatosis
- Purpura
- Graves' Disease (Thyroid Autoimmune)

Patient History
- cutaneous lupus
- osteoporosis
- hyperlipidemia
- frequent UTI
- hypothyroidism

Medications
- Alendronate
- pravastatin
- levothyroxine
- hydroxychloroquine

Findings
- urine dipstick:
- leukocyte esterase
- supine 120/80 mm HG
- heart rate: 88 bpm
- urine culture: E. Coli

Diagnosis Models:
- Renal Failure
- UTI
- Diabetes
- Influenza
- Hypokalemia
- Esophagitis
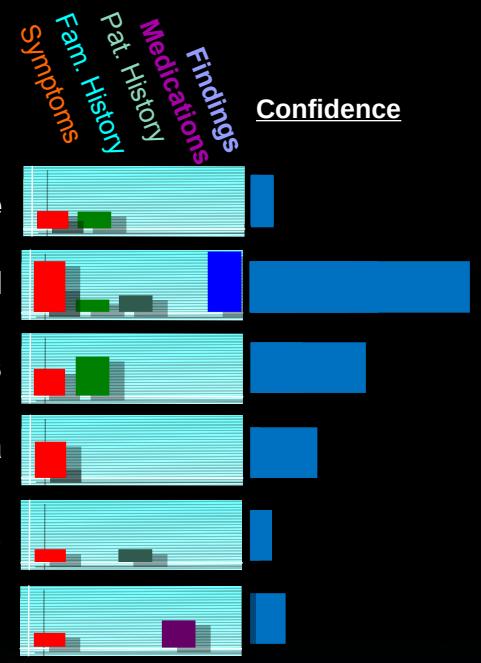
Most Confident Diagnosis: Esophagitis
- Extract Medications
- Use database of drug side-effects
- Together, multiple diagnoses may best explain symptoms
- Extract Findings: Confirms that UTI was present

Watson enables three classes of cognitive services

Ask

- Leverage vast amounts of data
- Ask questions for greater insights
- Natural language inquiries
- e.g. - Next generation Chat
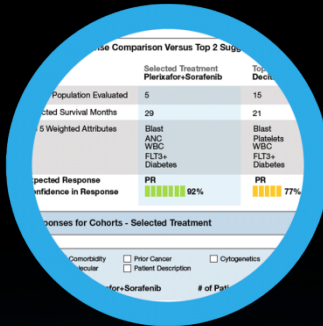
Discover

- Find the rationale for given answers
- Prompt for inputs to yield improved responses
- Inspire considerations of new ideas
- e.g. - Next generation Search → Discovery

Decide

- Ingest and analyze domain sources, info models
- Generate evidence based decisions with confidence
- Learn with new outcomes and actions
- e.g. - Next generation Apps → Probabilistic Apps

Putting Watson to work in healthcare

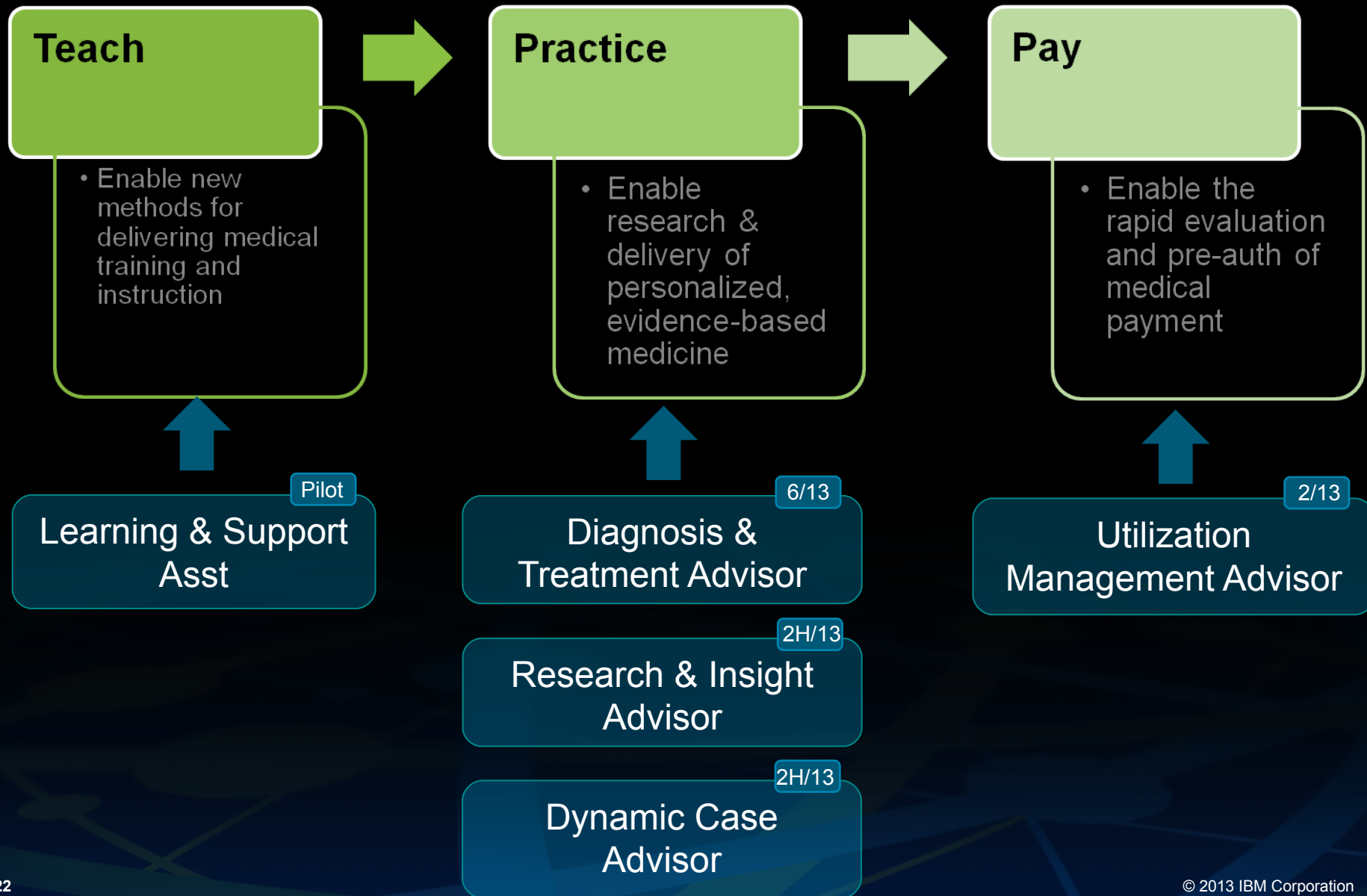| Teach		Practice		Pay
• Enable new methods for delivering medical training and instruction	→	• Enable research & delivery of personalized, evidence-based medicine	→	• Enable the rapid evaluation and pre-auth of medical payment

Pilot

Learning & Support Asst

6/13

Diagnosis & Treatment Advisor

2H/13

Research & Insight Advisor

2H/13

Dynamic Case Advisor

2/13

Utilization Management Advisor

Smarter Care – We have the capabilities to deliver

Research & Insights	Diagnose & Treat	Plan, Engage & Coordinate	Measure & Support
- Analyze population segments - Obtain insights from a large amount of structured/un-structured data to define new treatments - Clinical trials - **Stakeholders** – Researchers, Specialists	- Leverage analytics and insights to diagnose specific conditions - Define treatment plans - **Stakeholders** – PCP's, Specialists, Nurse Practitioners	- Define a care plan, assign multi-disciplinary teams to manage care - Coordinate care amongst stakeholders - **Stakeholders** – Care coordinator, nurses, PCPs, Specialists, 3rd party service providers	- Define an outcome plan - Continuously assess progress towards delivering results - **Stakeholders** – Care coordinators

ECM   **Research**

Watson Research Advisor (RA), Watson Dynamic Case Advisor, ECM LanguageWare, ECM ICPA, ICDA

CÚRAM SOFTWARE an IBM Company   **ECM**

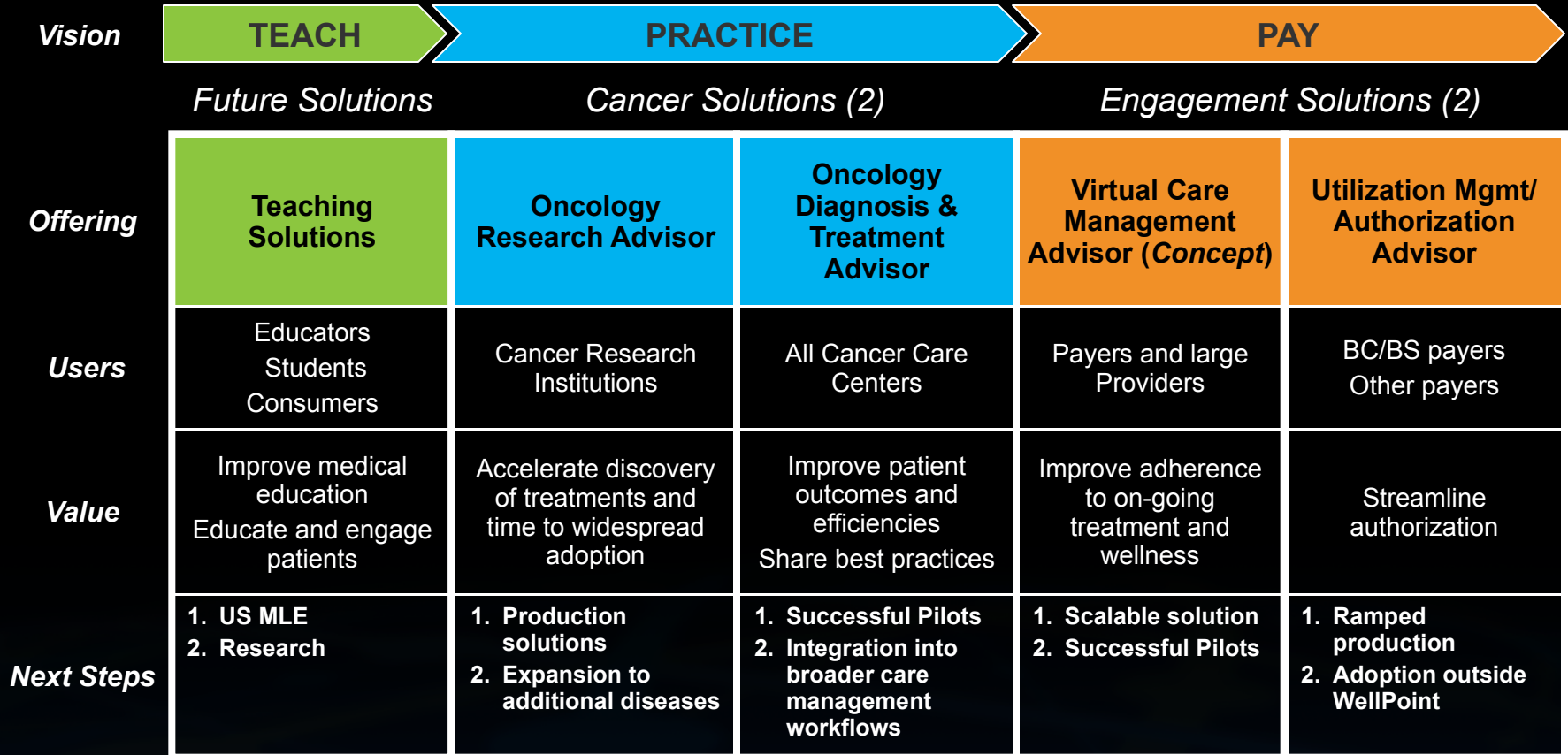*Supporting*   SPSS AN IBM COMPANY   COGNOS   Initiate

- Extends health with social determinants to provide a holistic view of an individual
- Focuses on high cost, high need and high risk individuals
- Enables collaboration in a complex ecosystem aimed at improving outcomes

23

Topics

- What is Big Data and why does Big Data Analytics matter?

- How does IBM Watson leverage Big Data Analytics in Healthcare?

- What are some of the early results from Watson customers?

Watson for healthcare 2013

Vision	TEACH	PRACTICE		PAY	
	Future Solutions	*Cancer Solutions (2)*		*Engagement Solutions (2)*	
Offering	Teaching Solutions	Oncology Research Advisor	Oncology Diagnosis & Treatment Advisor	Virtual Care Management Advisor (*Concept*)	Utilization Mgmt/ Authorization Advisor
Users	Educators Students Consumers	Cancer Research Institutions	All Cancer Care Centers	Payers and large Providers	BC/BS payers Other payers
Value	Improve medical education Educate and engage patients	Accelerate discovery of treatments and time to widespread adoption	Improve patient outcomes and efficiencies Share best practices	Improve adherence to on-going treatment and wellness	Streamline authorization
Next Steps	1. US MLE 2. Research	1. **Production solutions** 2. **Expansion to additional diseases**	1. **Successful Pilots** 2. **Integration into broader care management workflows**	1. **Scalable solution** 2. **Successful Pilots**	1. **Ramped production** 2. **Adoption outside WellPoint**

Today Watson is helping payers streamline pre-approvals

THEN

Written authorization / subjective assessment / slow response

NEXT

Fact based – data driven analysis / systematic review & recommendation / timely response

NOW

Manual review / pre-approval / code based procedures / guideline driven

NEJM medical concept annotations

Diseases   **Symptoms**   **Medications**   **Modifiers**

Relations
causeOf
modifierOf
negationOf
partOf
remedyOf
resultOf

1 Chamarthi, Bindu; Morris, Charles A.; Kaiser, Ursula B.; Katz, Joel T.; Loscalzo, Joseph

2 Stalking the Diagnosis

3 362/9/834

4 http://content.nejm.org/cgi/content/full/362/9/834</citation_fulltext_html_url>

5 A 58-year-old woman presented to her primary care physician after several days of dizziness, anorexia, dry mouth, increased thirst, and frequent urination. She had also had a fever and reported that food would "get stuck" when she was swallowing. She reported no pain in her abdomen, back, or flank and no cough, shortness of breath, diarrhea, or dysuria. Her history was notable for cutaneous lupus, hyperlipidemia, osteoporosis, frequent urinary tract infections, three uncomplicated cesarean sections, a left oophorectomy for a benign cyst, and primary hypothyroidism, which had been diagnosed a year earlier. Her medications were levothyroxine, hydroxychloroquine, pravastatin, and alendronate. She lived with her husband and had three healthy adult children. She had a 20-pack-year history of smoking but had quit 3 weeks before presentation. She reported no alcohol or drug abuse and no exposure to tuberculosis. Her family history included oral and bladder cancer in her mother, Graves' disease in two sisters, hemochromatosis in one sister, and idiopathic thrombocytopenic purpura in one sister.

Entity Types / Roles
FAMILY-DISEASE
FAMILY-SUBSTANCE-ABUSE
FINDING-BLOODPRESSURE
FINDING-GENERIC
FINDING-HEARTRATE
FINDING-HEIGHT
FINDING-OXYGEN-SATURATIO
FINDING-RESPIRATORYRATE
FINDING-TEMPERATURE
FINDING-WEIGHT
MODIFIER-ANATOMY
MODIFIER-GENERIC
MODIFIER-NEGATION
MODIFIER-TIME
PATIENT-ACTIVITY-EVENT
PATIENT-AGE
PATIENT-ALLERGY
PATIENT-FEMALE
PATIENT-HAZARD-EXPOSURE
PATIENT-HEALTHSTATE
PATIENT-LOCATION
PATIENT-MALE
PATIENT-NAME
PATIENT-OCCUPATION

Watson is helping improve how medicine is practiced

THEN

Paper based files / limited knowledge sharing / experience based decisions

NEXT

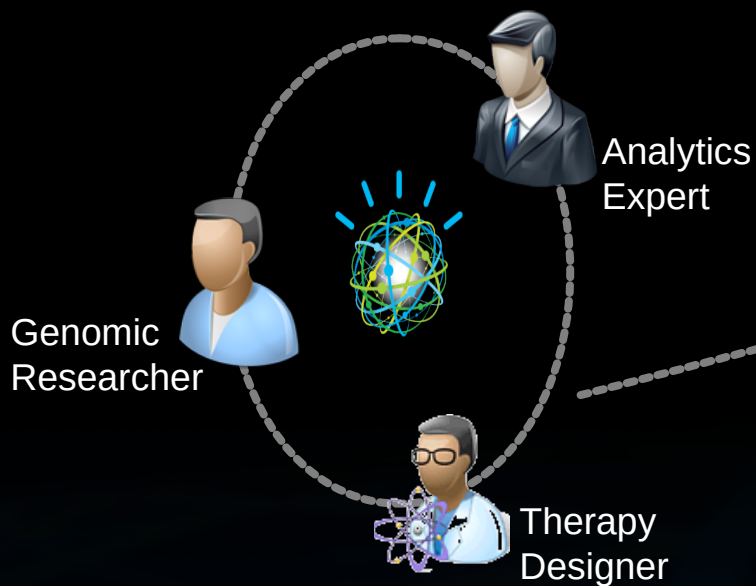Real time insights / full transparency / confidence weighted actions

NOW

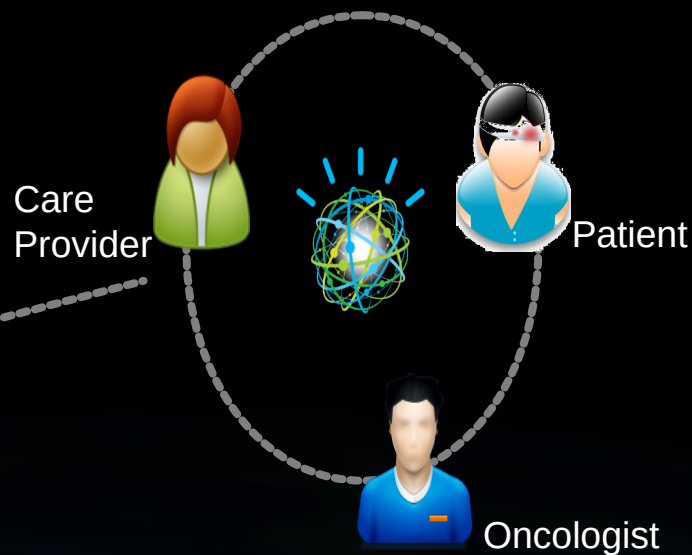Electronic medical records / information overload / 55% medicine evidence based

Watson goes to work to help researchers as well

Accelerate Time to Clinical Insights

Analytics Expert

Genomic Researcher

Therapy Designer

Medical Research

Improve Decisions and Outcomes

Care Provider

Patient

Oncologist

Medical Practice & Payments

Ultimate Goal: Become the Most Essential Company

IBM Oncology Diagnosis and Treatment Advisor

Shows how Watson can assist an Oncologist by:

- Synthesizing disparate data – patient records, clinician notes, test results, pathology reports, etc.

- Identifying missing pieces of data recommending tests with complete transparency

- Suggesting personalized, confidence-weighted, evidence-based options to improve quality of care and patient experience

Treatment Plan	Confidence	Patient Preferences Match
Treatment plan 1 Systemic Chemo: Cisplatin, Pemetrexed, Bevacizumab	95%	Acceptable match with patient preferences
Treatment plan 2 Systemic Chemo: Carboplatin, Paclitaxel, Bevacizumab	45%	Unnaceptable match with patient preferences
Treatment plan 3 Systemic Chemo: Erlotinib	8%	Preferred match with patient preferences

DEMONSTRATION

Watson Research & Insight Advisor

Provides research oncologists the ability to:

- Dynamically build a patient cohort

- Suggest confidence-weighted treatment options for that cohort

- Generate insights on cohort trends from clinical, medical, and patient data

- Accelerate treatment identification, publication and distribution

DEMONSTRATION

How it Works: Cloud Delivery and Outcome Based Pricing

Dynamic Capacity

Automate and control
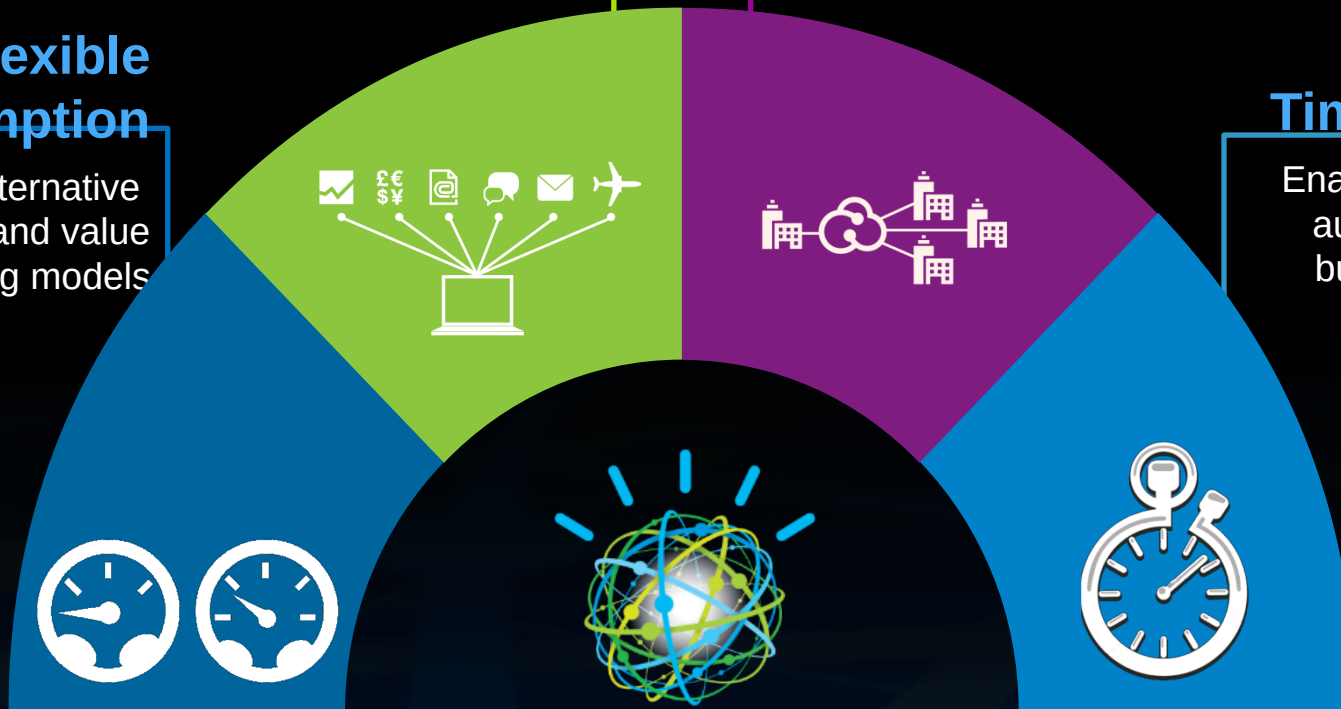service provisioning

Hybrid Delivery

Extend & integrate
on-premise solution
with cloud offering
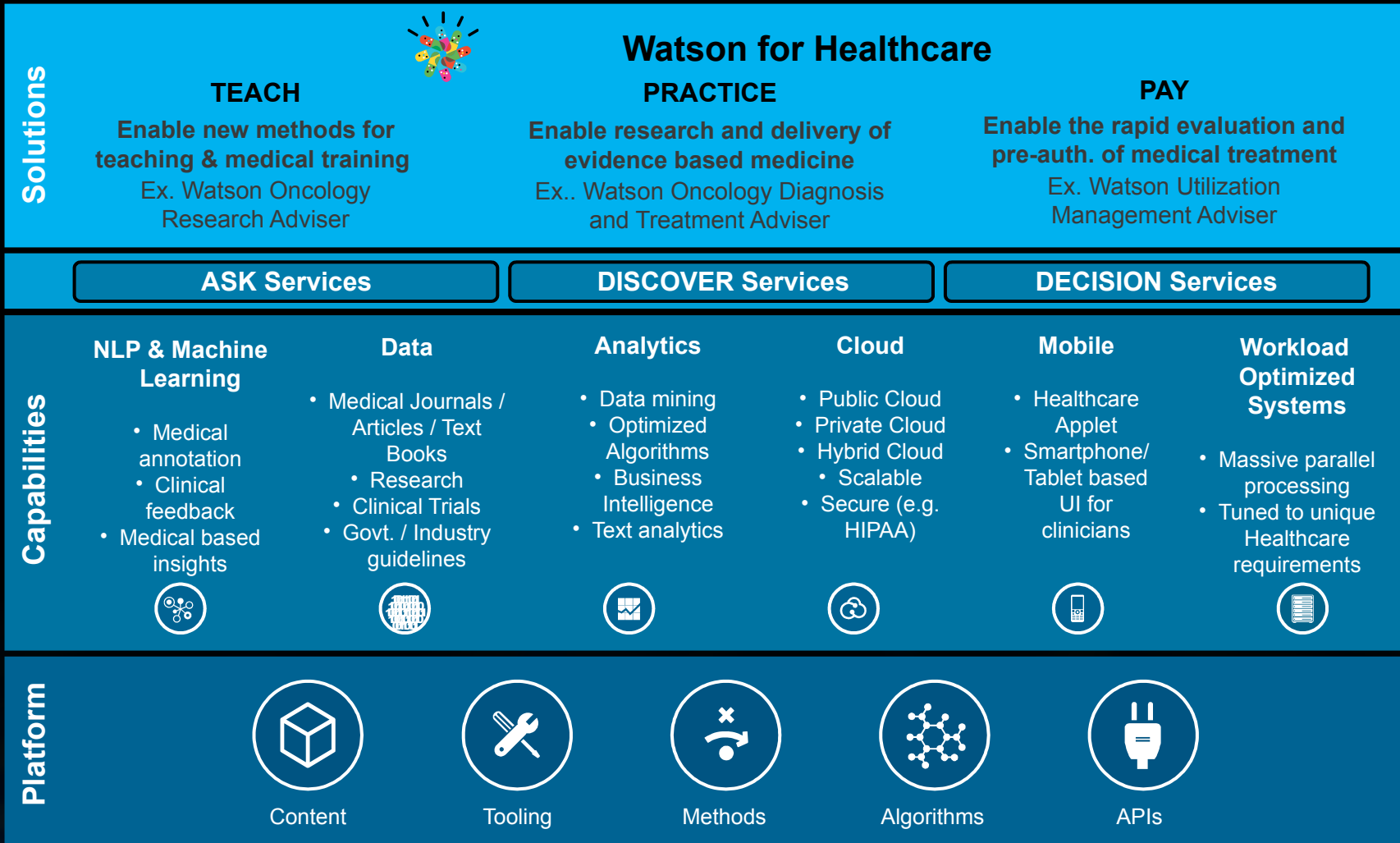
Flexible Consumption

Support alternative
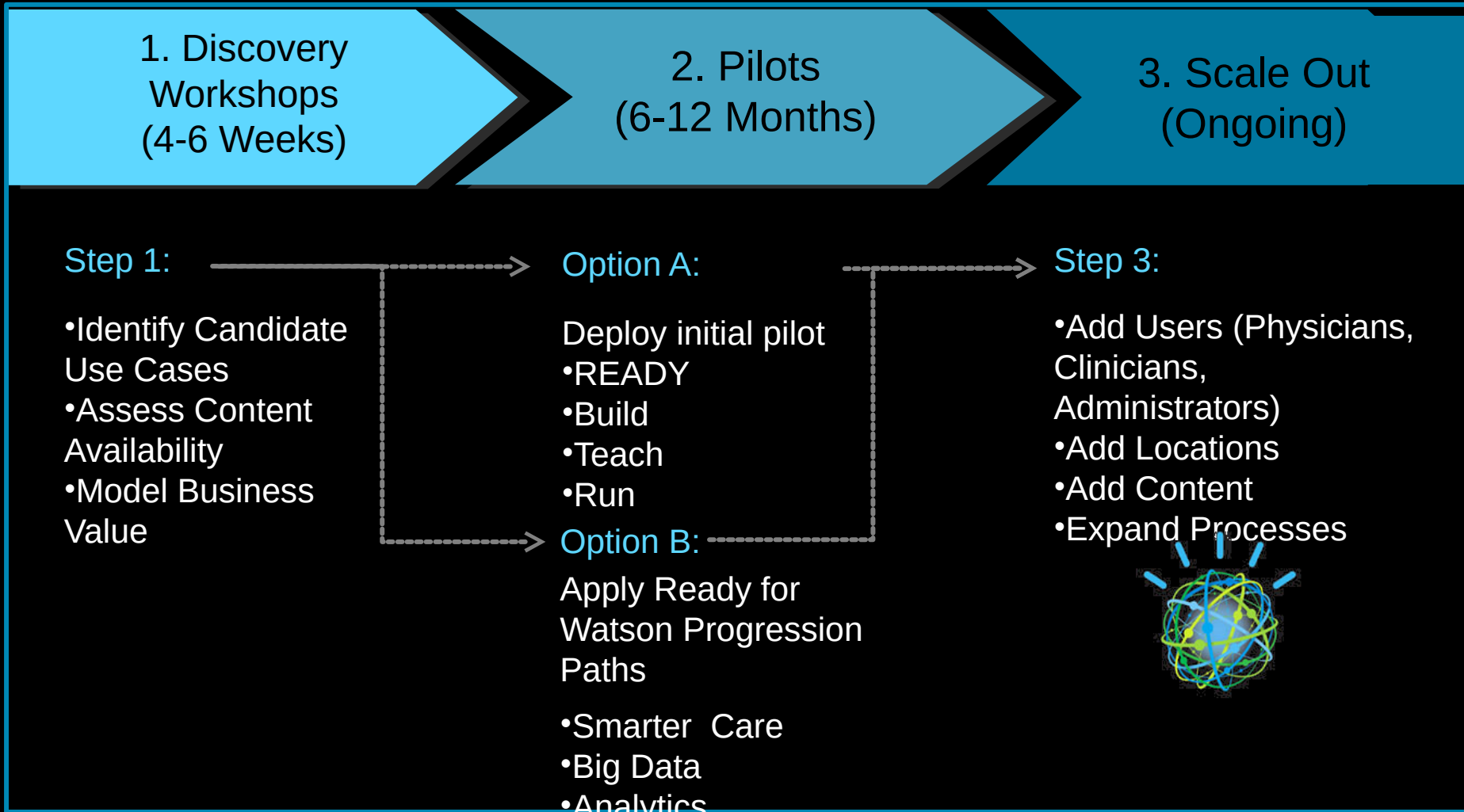delivery and value
pricing models

Time to Value

Enable incremental
automation and
business agility

How it Works: Watson healthcare architecture

Solutions

Watson for Healthcare

TEACH	PRACTICE	PAY
Enable new methods for teaching & medical training	**Enable research and delivery of evidence based medicine**	**Enable the rapid evaluation and pre-auth. of medical treatment**
Ex. Watson Oncology Research Adviser	Ex.. Watson Oncology Diagnosis and Treatment Adviser	Ex. Watson Utilization Management Adviser

ASK Services	DISCOVER Services	DECISION Services

Capabilities

NLP & Machine Learning	Data	Analytics	Cloud	Mobile	Workload Optimized Systems
• Medical annotation • Clinical feedback • Medical based insights	• Medical Journals / Articles / Text Books • Research • Clinical Trials • Govt. / Industry guidelines	• Data mining • Optimized Algorithms • Business Intelligence • Text analytics	• Public Cloud • Private Cloud • Hybrid Cloud • Scalable • Secure (e.g. HIPAA)	• Healthcare Applet • Smartphone/ Tablet based UI for clinicians	• Massive parallel processing • Tuned to unique Healthcare requirements

Platform

Content	Tooling	Methods	Algorithms	APIs

Getting Started with Watson

1. Discovery Workshops (4-6 Weeks)	2. Pilots (6-12 Months)	3. Scale Out (Ongoing)

Step 1:

- Identify Candidate Use Cases
- Assess Content Availability
- Model Business Value

Option A:

Deploy initial pilot
- READY
- Build
- Teach
- Run

Option B:

Apply Ready for Watson Progression Paths

- Smarter Care
- Big Data
- Analytics

Step 3:

- Add Users (Physicians, Clinicians, Administrators)
- Add Locations
- Add Content
- Expand Processes

How it Works: Watson Products and Infrastructure
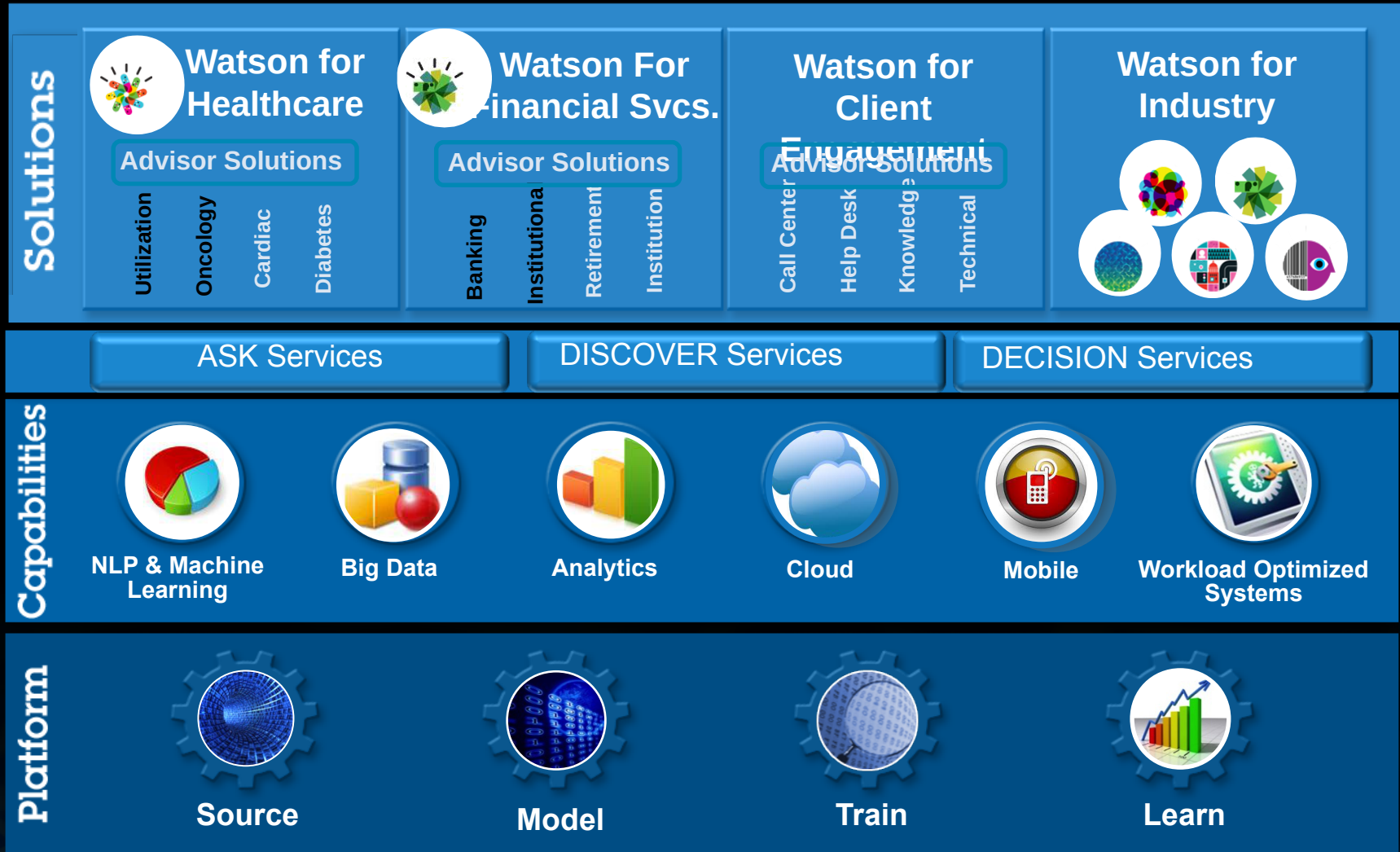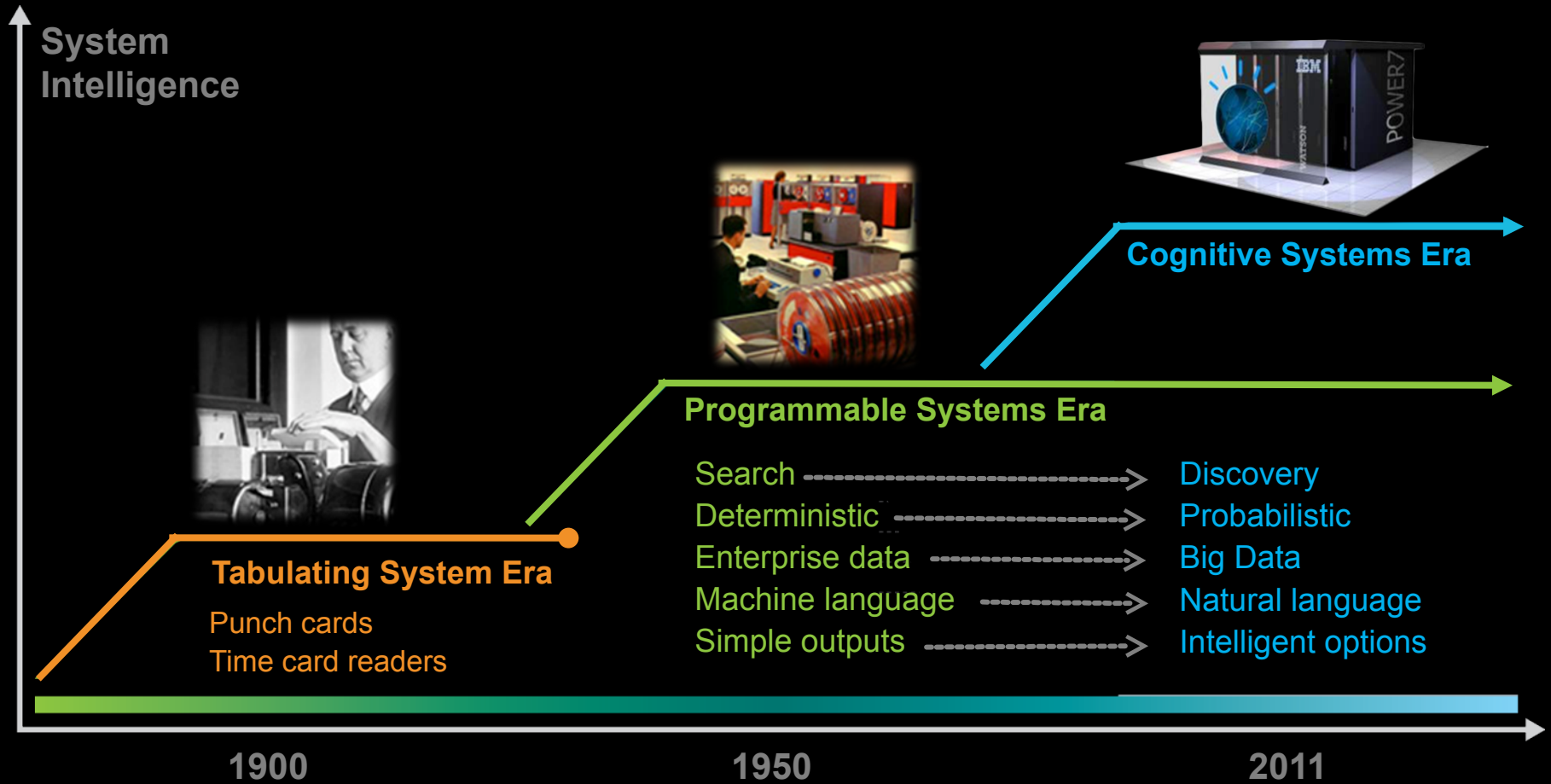
Thank you!
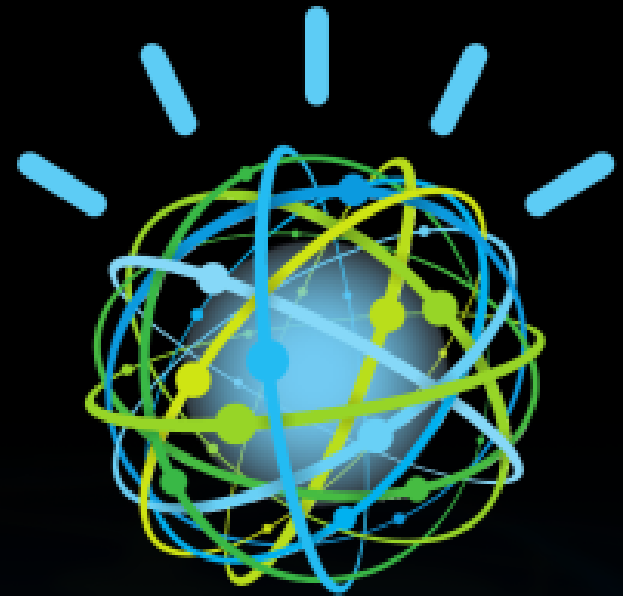
Please note

IBM's statements regarding its plans, directions, and intent are subject to change or withdrawal without notice at IBM's sole discretion.

Information regarding potential future products is intended to outline our general product direction and it should not be relied on in making a purchasing decision.

The information mentioned regarding potential future products is not a commitment, promise, or legal obligation to deliver any material, code or functionality. Information about potential future products may not be incorporated into any contract. The development, release, and timing of any future features or functionality described for our products remains at our sole discretion.

Performance is based on measurements and projections using standard IBM benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon many factors, including considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve results similar to those stated here.