Bringing Data to Action

http://bizbench.com/wisdom-pyramid/

1001 Free Things You Can Do With Your Data to Achieve Analytics Mastery

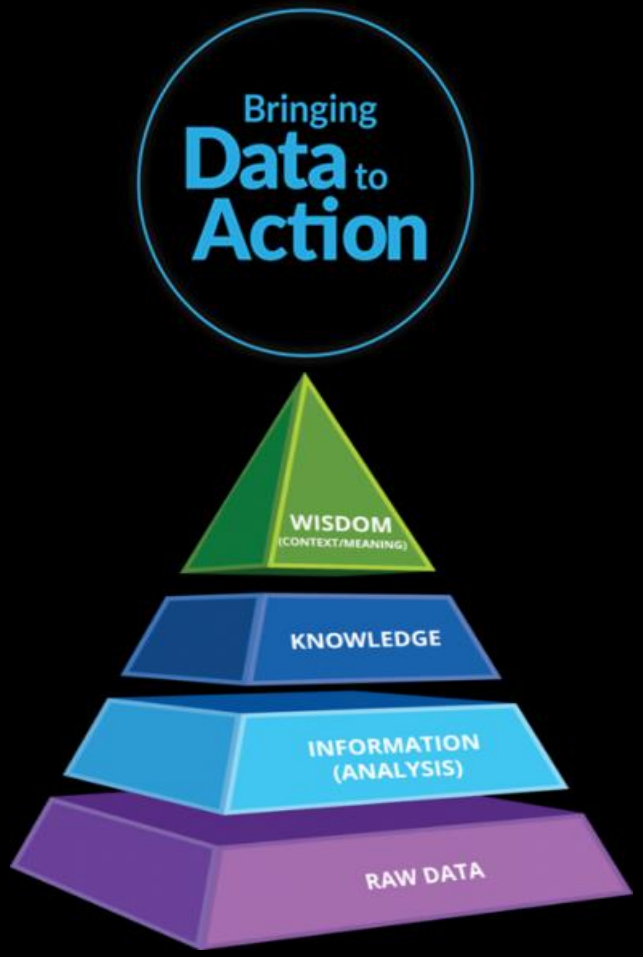
Kirk Borne

@KirkDBorne

These slides here: http://www.kirkborne.net/datadotworld2021/

OUTLINE

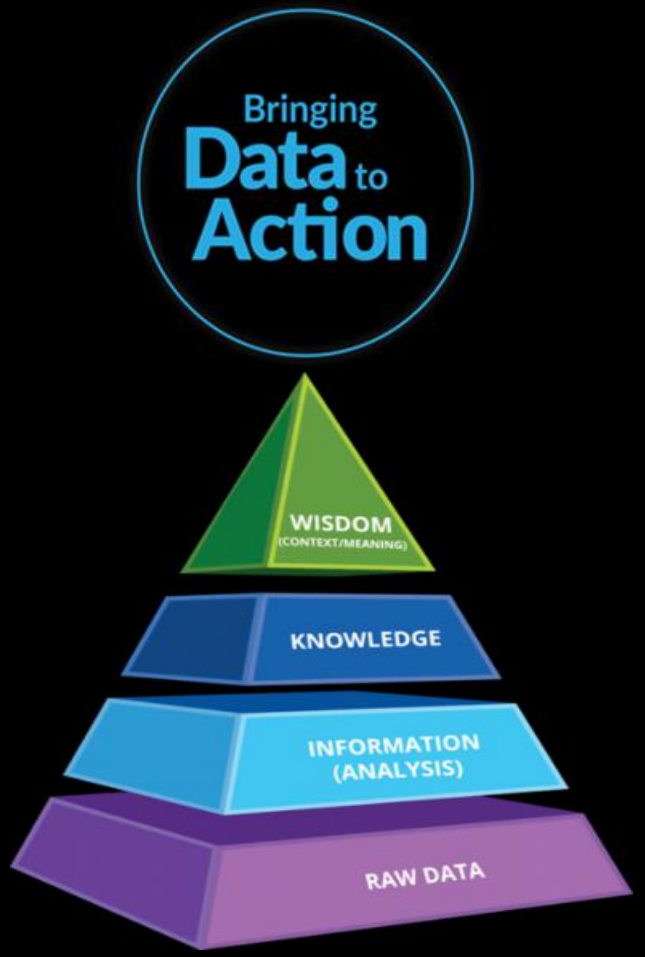
- Preliminaries
- Having the “Talk”
- Self-Driving Enterprise
- 1001 Analytics Things to Do



http://bizbench.com/wisdom-pyramid/

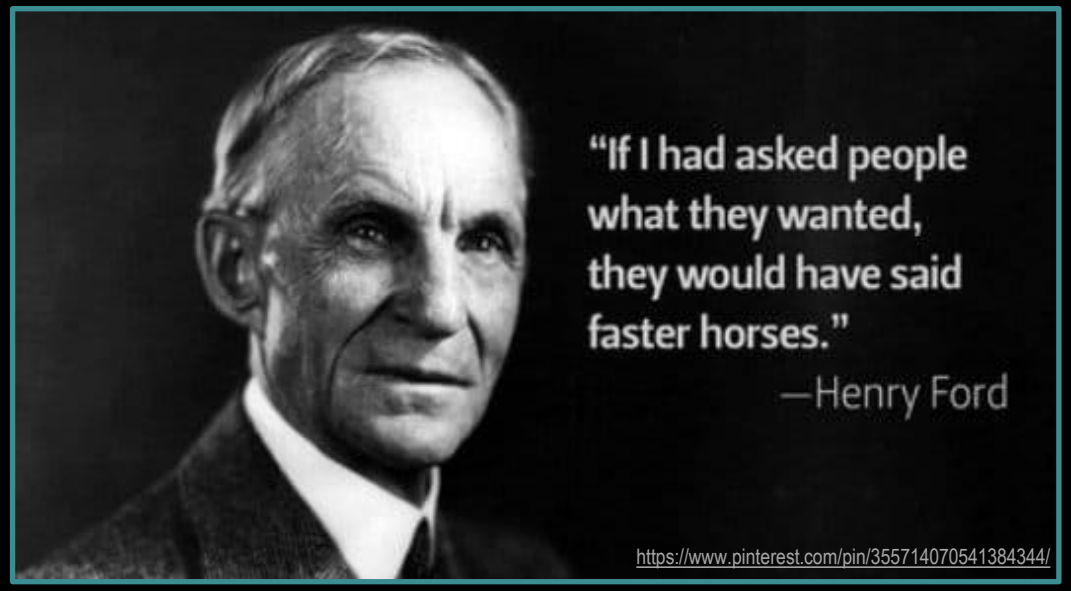
OUTLINE

- **Preliminaries**
- Having the “Talk”
- Self-Driving Enterprise
- 1001 Analytics Things to Do



http://bizbench.com/wisdom-pyramid/

Business Transformation – then and now

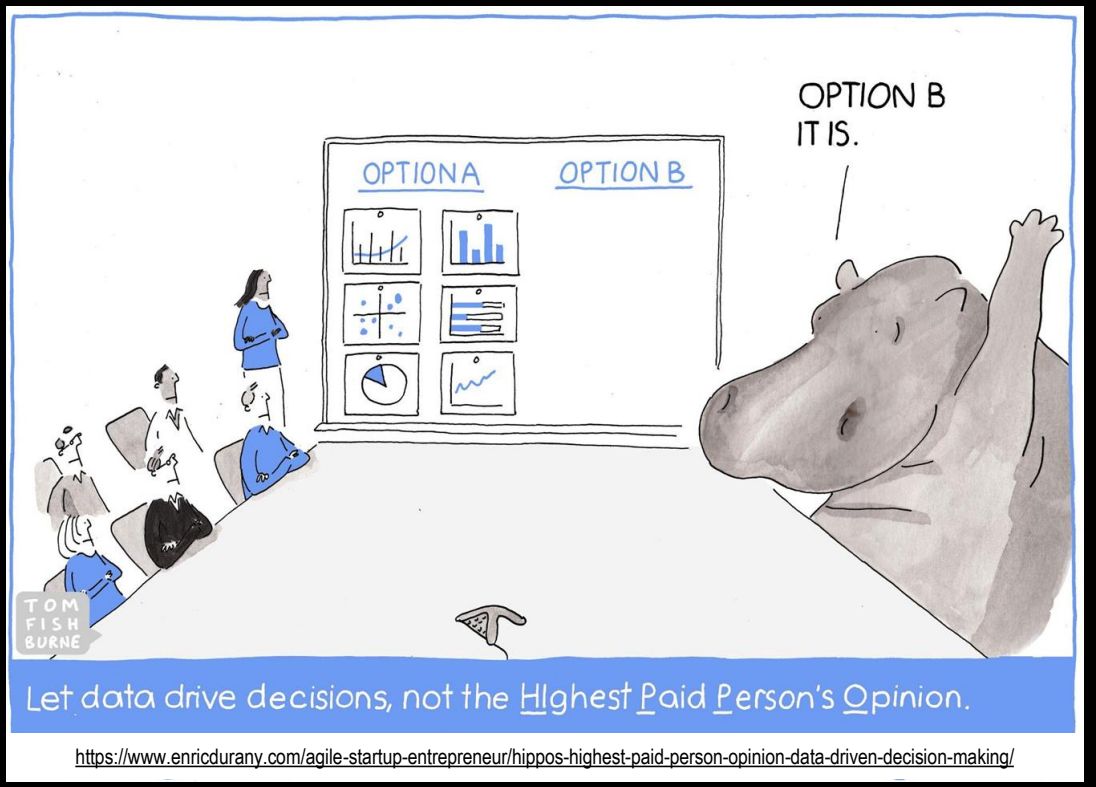


"If CIOs invested in machine learning three years ago, they would have wasted their money. But if they wait another three years, they will never catch up."

- Dan Olley, CTO of Elsevier

https://www.cio.com/article/3061713/leadership-management/why-its-time-for-cios-to-invest-in-machine-learning.html

Beware the HiPPO in the room!



https://www.enricdurany.com/agile-startup-entrepreneur/hippos-highest-paid-person-opinion-data-driven-decision-making/

Let the data drive the decisions!

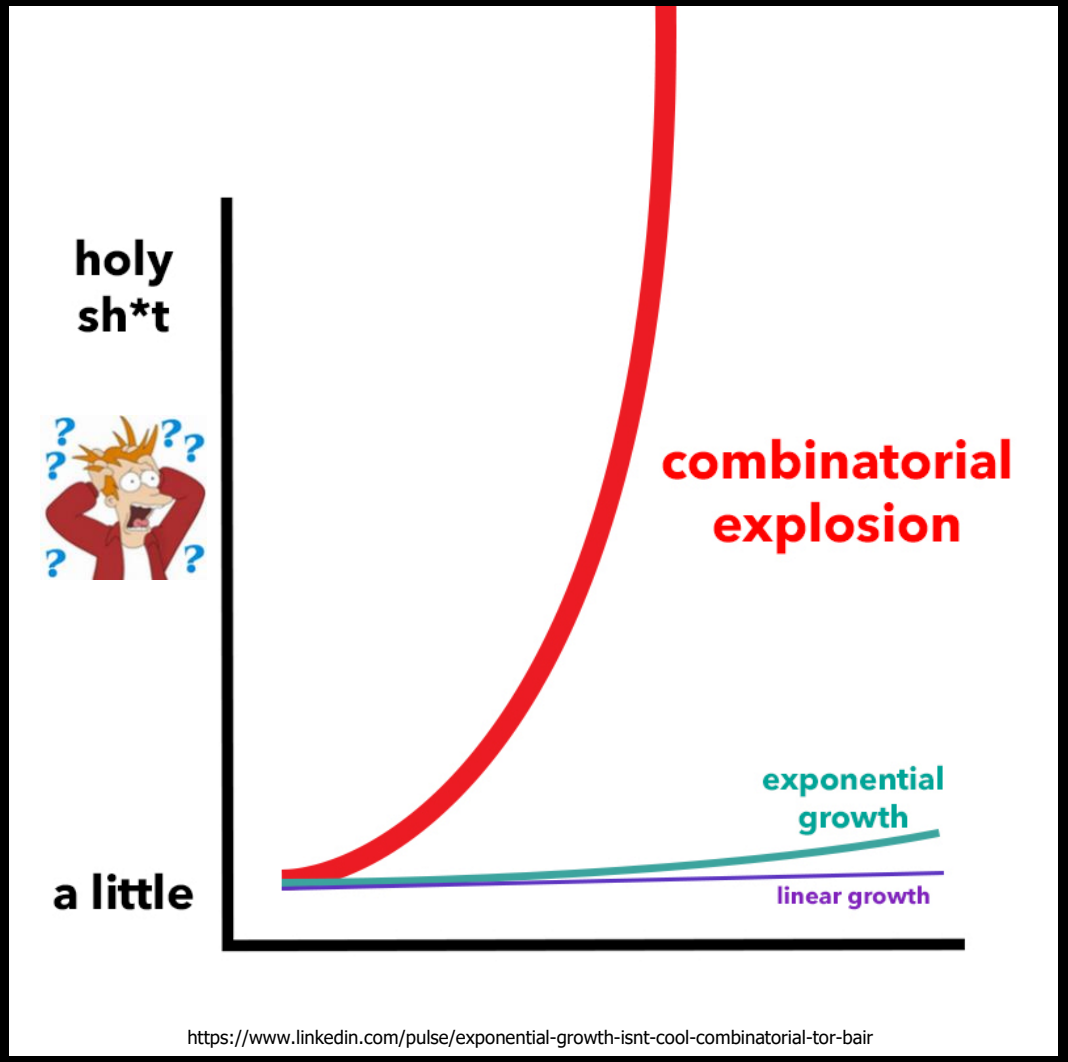
Let us rewind back to the beginning…

Ever since we first explored our world…

…we have asked questions about everything around us.

https://allthatsinteresting.com/worlds-most-incredible-rock-formations

So, we have collected evidence (data) to answer our questions, which leads to more questions, which leads to more data collection, which leads to more questions, …, which leads to **BIG DATA!**



https://www.linkedin.com/pulse/exponential-growth-isnt-cool-combinatorial-tor-bair

Knowledge is about connecting the dots.

@KirkDBorne

$y \sim x! \approx x \wedge x$

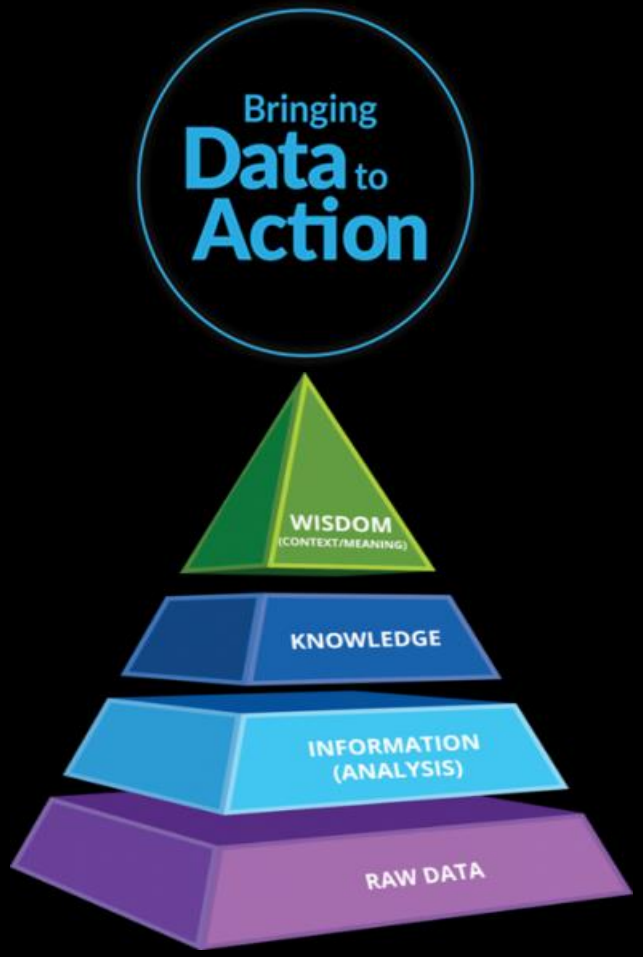
→ Combinatorial Growth!
(all possible interconnections, linkages, and interactions)

$y \sim 2 \wedge x$ (exponential growth)

$y \sim 2 * x$ (linear growth)

OUTLINE

- Preliminaries
- **Having the "Talk"**
- Self-Driving Enterprise
- 1001 Analytics Things to Do



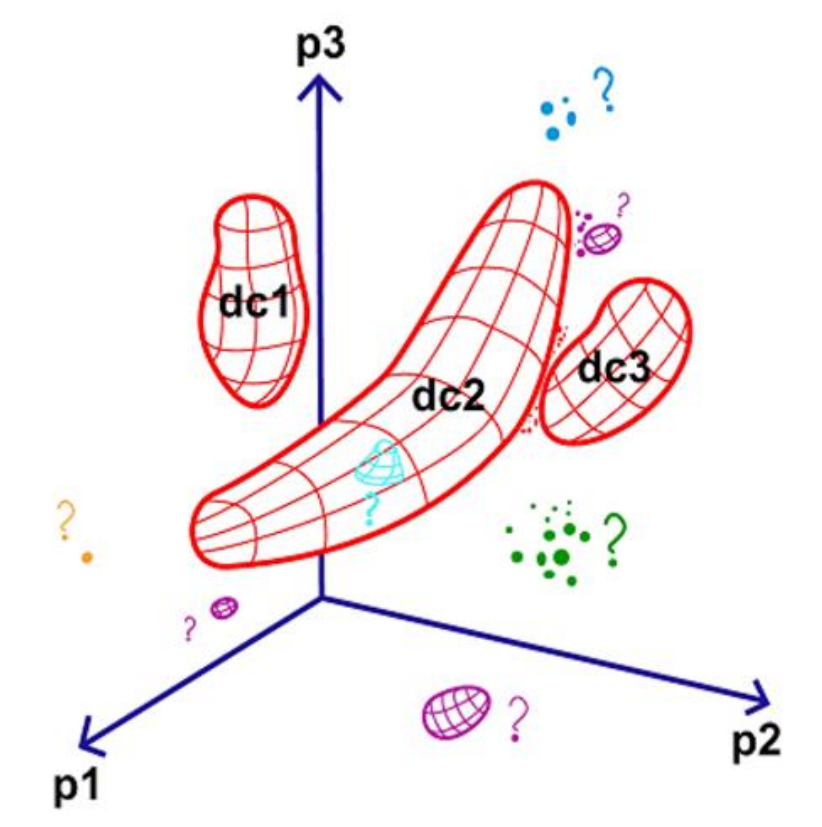
http://bizbench.com/wisdom-pyramid/

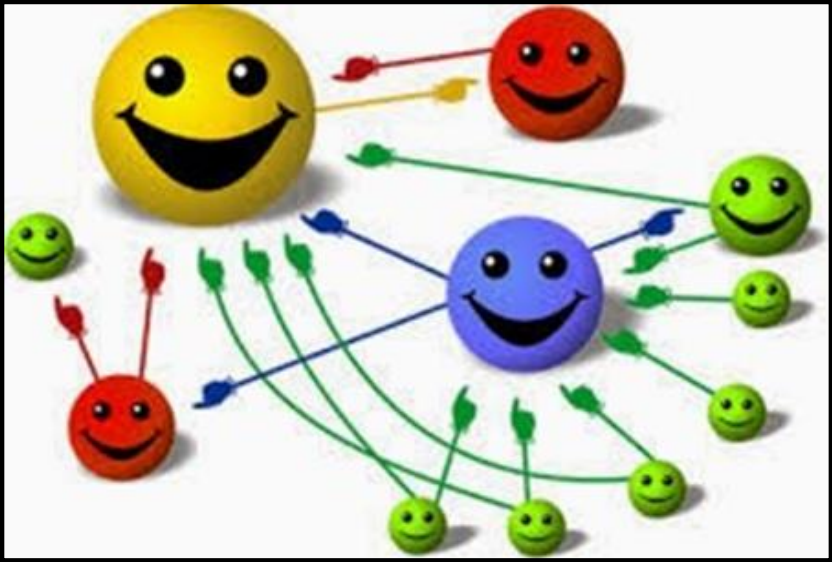
Having the talk – How ML takes Data to Insights to Business Value

Knowing the "Easy Button" for Identifying Insights and Extracting Value from Data through Machine Learning

- Pattern Discovery (Detection)
  - D2D: Data-to-Discovery
- Pattern Recognition
  - D2D: Data-to-Decisions
- Pattern Exploration
  - D2D: Data-to-Dollars (innovation)
- Pattern Exploitation
  - D2B: Data-to-Best action / Best value

Insights Discovery



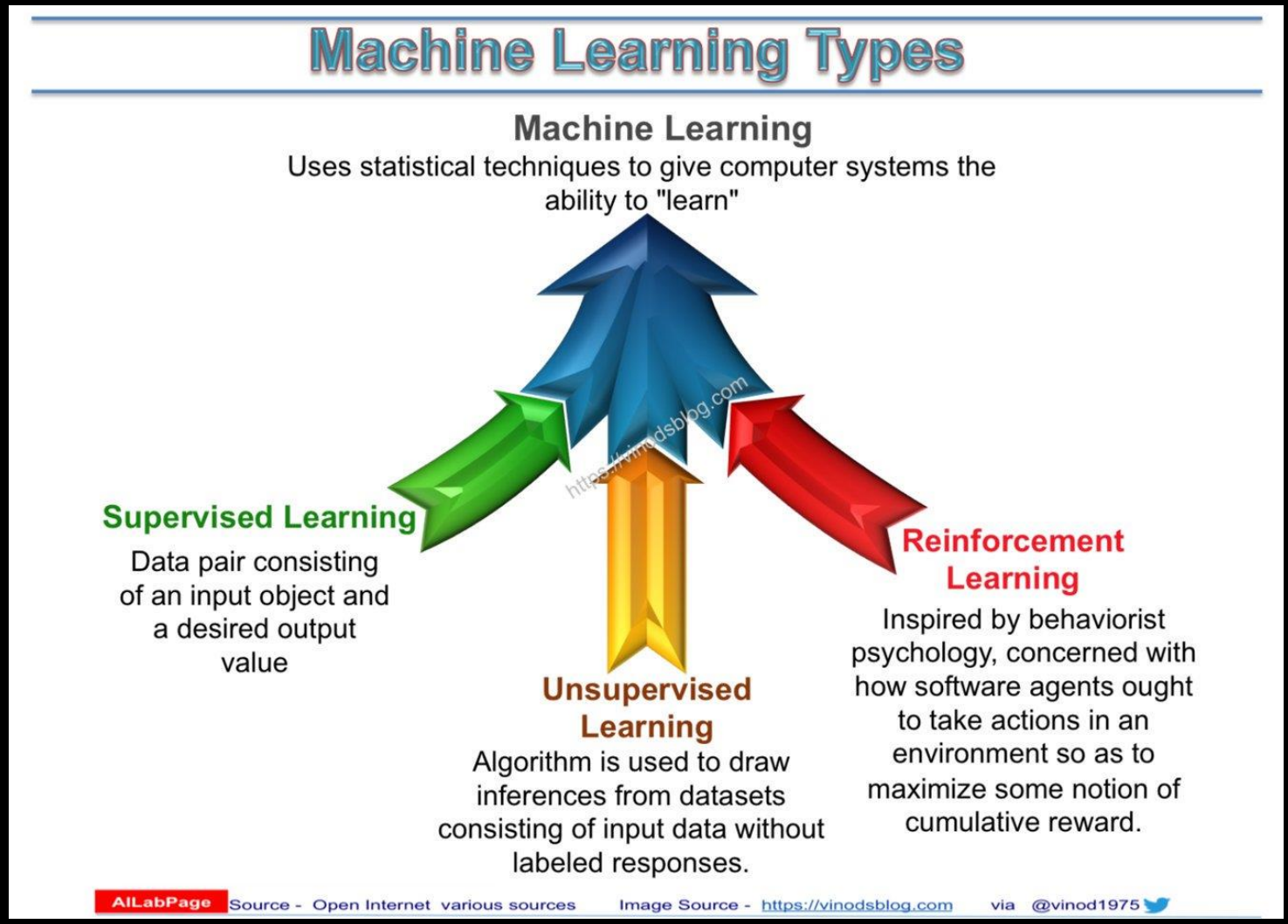


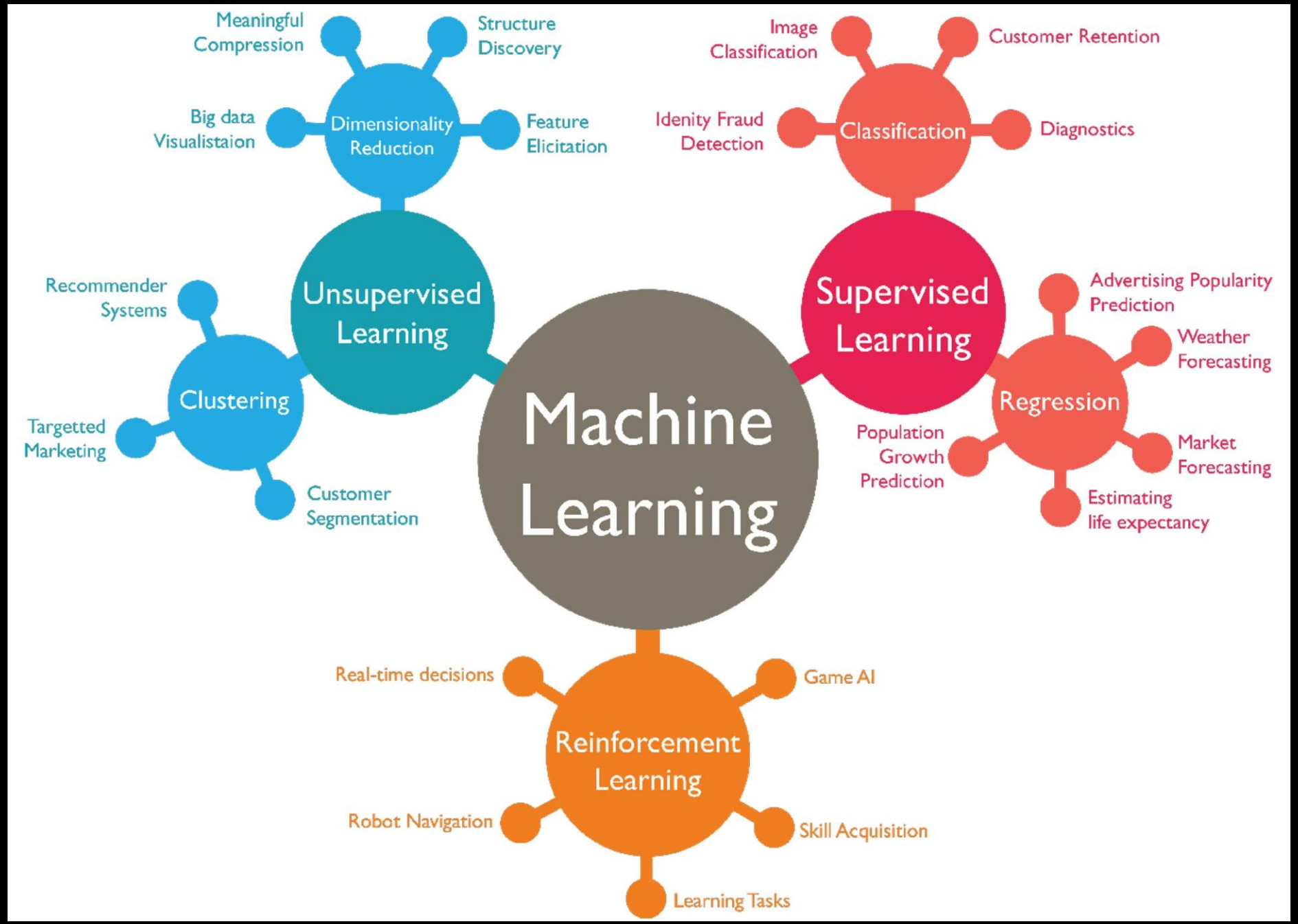
Data is not a 4-letter word, and neither is Machine Learning. These can be used by everyone for business insights:

1) **Class Discovery:** Find the categories (population segments & sub-segments) of things (objects, events, and behaviors) in your data. + Learn the rules that constrain the different class boundaries (that uniquely distinguish the segments).
2) **Correlation (Predictive and Prescriptive Power) Discovery:** (insights discovery) – Find trends, patterns, dependencies in data that reveal the governing principles, causal relations, and/or behavioral patterns of things (their "DNA").
3) **Outlier / Anomaly / Novelty / Surprise Discovery:** Find the new, surprising, unexpected one-in-a-[million / billion / trillion] object, event, or behavior in the data.
4) **Association (or Link) Discovery:** (Graph and Network Analytics) – Find both the usual and the unusual (interesting) data associations / links / connections across the entities in your domain. "Connects the dots that aren't connected."

Machine Learning =
Mathematical Algorithms that Learn from Experience

https://bit.ly/2jecUdE

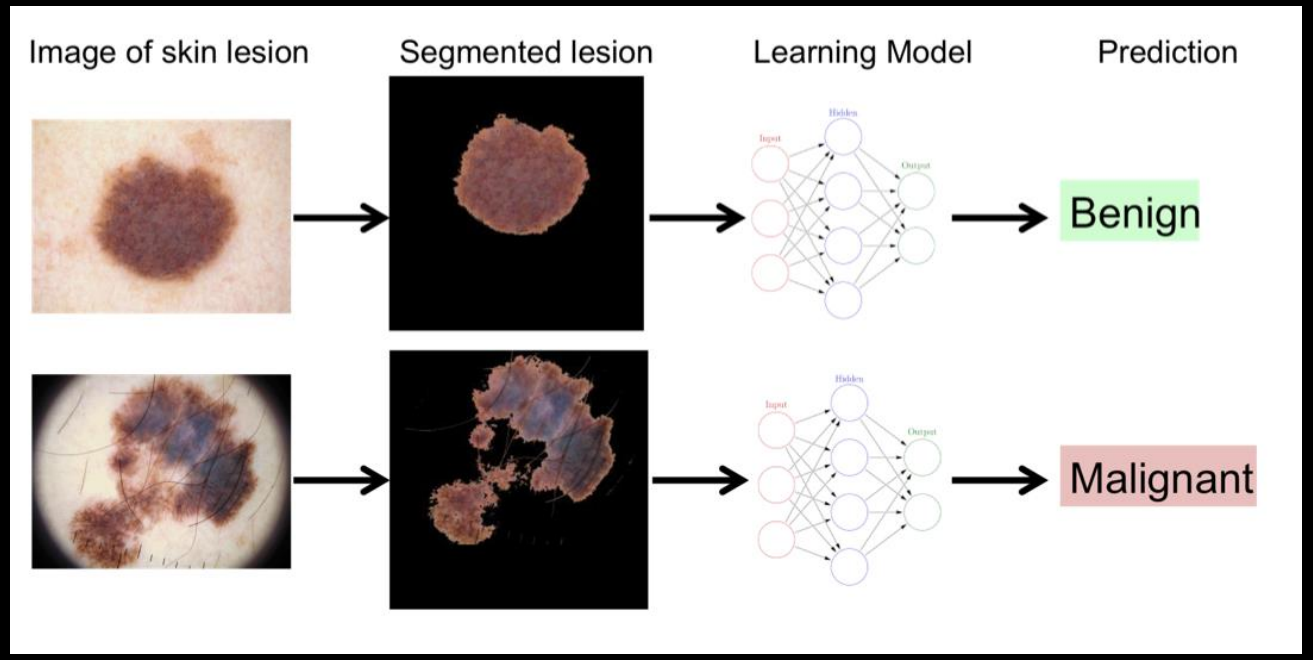




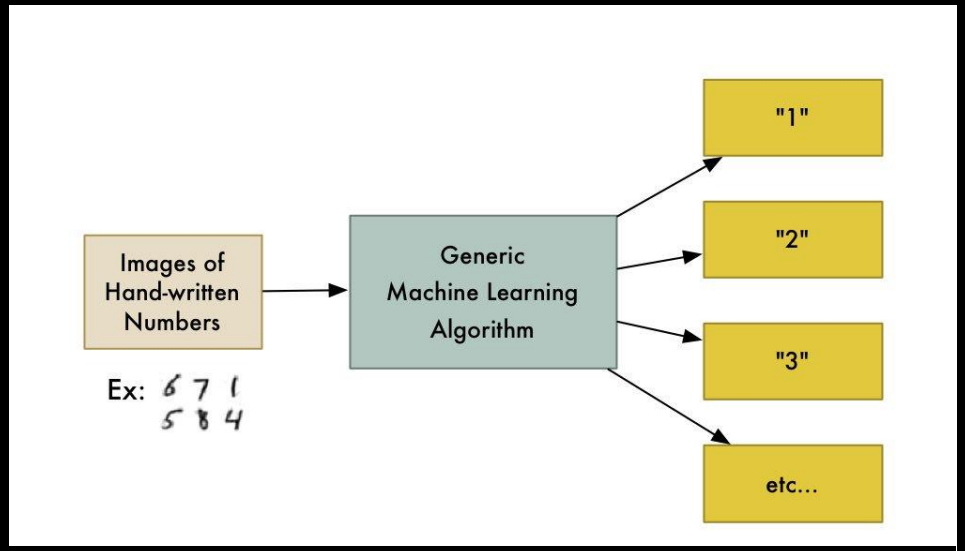
https://cloudsmallbusinessservice.com/blog/5-best-machine-learning-frameworks-of-2020.html

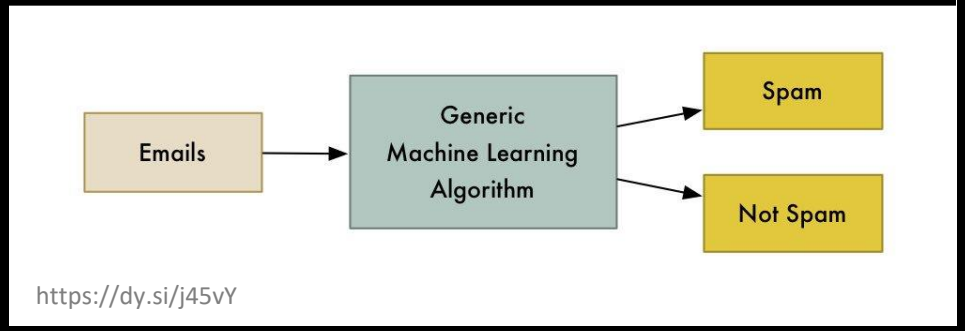
Machine Learning = Algorithms that Learn from Experience <=> **Models that learn characteristic patterns in data to get things right!**

Examples: (1) Digit detection algorithm (used in automated Postal Code readers). (2) Email Spam detection algorithm (used for Spam filtering). (3) Cancer detection algorithm (used in medical imaging diagnosis). (4) Credit Card Fraud Alerts. (5) Siri, Alexa chatbots (Conversational AI). (6) Online Product Recommendation Engines.



https://web.stanford.edu/~kalouche/cs229.html





https://dy.si/j45vY

5 Dimensions of Analytics Outcomes in Data-intensive Business Applications

1) Descriptive Analytics

- **Hindsight** (What happened?)
- Asks the required questions.

2) Diagnostic Analytics

- **Oversight** (Real-time / What is happening? Why did it happen?)

3) Predictive Analytics

- **Foresight** (What will happen?)

4) Prescriptive Analytics

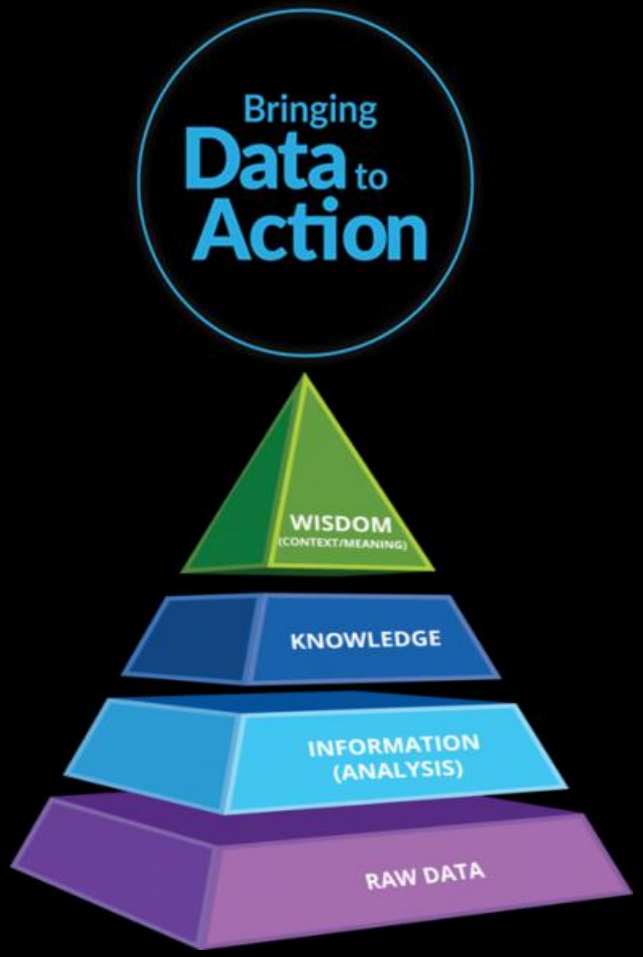
- **Insight** (How can we optimize what happens?) (Follow the dots!)

5) Cognitive Analytics

- **Right Sight** (the 360 view; when data informs the right decision and the right action right now, in the right context)
- Moves beyond simply providing answers, to **generating new questions from the data.**

OUTLINE

- Preliminaries
- Having the "Talk"
- **Self-Driving Enterprise**
- 1001 Analytics Things to Do



http://bizbench.com/wisdom-pyramid/

What is a Self-Driving Enterprise?

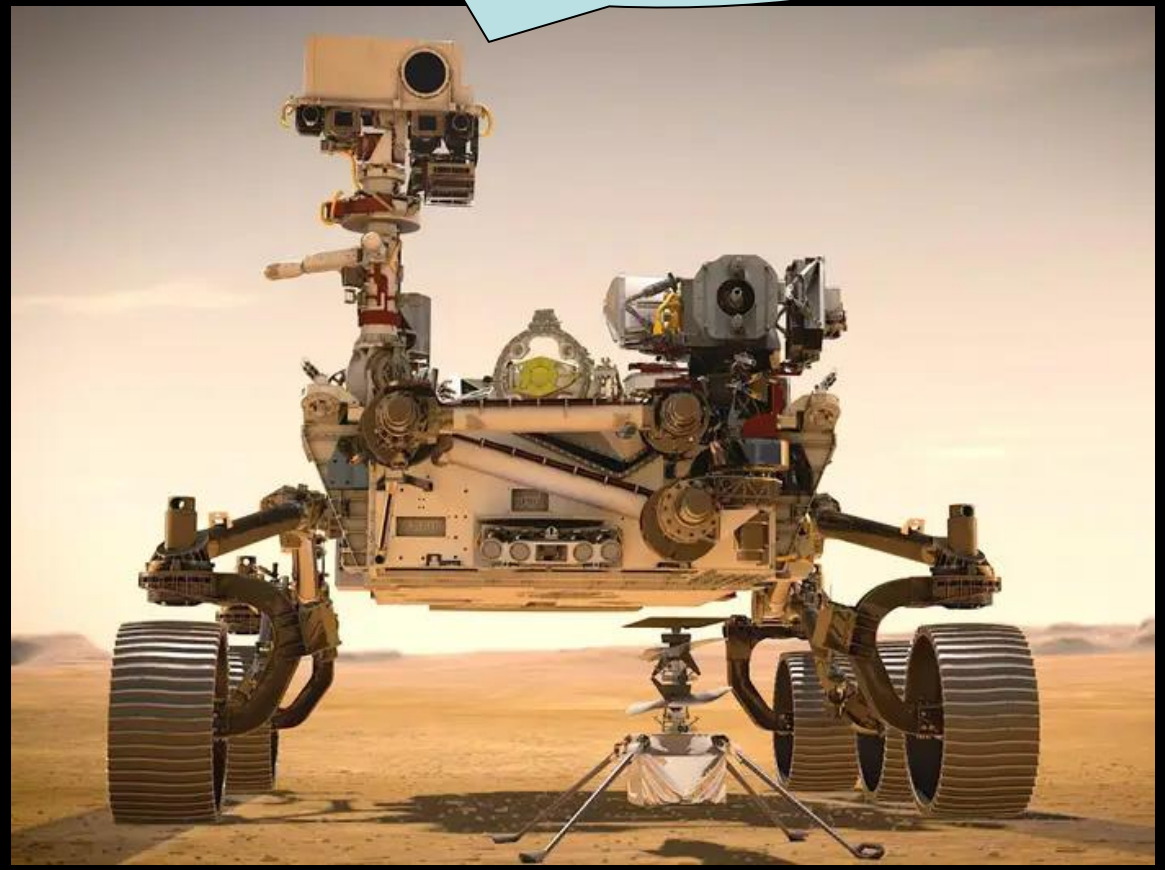
The self-driving (autonomous) enterprise is any entity that learns, gains actionable insights, discovers next-best action, innovates, and creates value from streaming big data and edge analytics via ubiquitous internal and external (contextual) sensors and data sources.

Mars Rovers :

- Metaphor for your use case!
- Smart data gatherer
- Actionable intelligent agent
- Autonomous decision system for the data-informed journey.

http://bizbench.com/wisdom-pyramid/

"Hello! I'm Percy."
(Perseverance Rover)



https://www.businessinsider.com/nasa-first-interplanetary-helicopter-in-perseverance-mars-rover-belly-2020-7

Taking Data to Information to Knowledge to Understanding to Action

Smart Sensors & Sentinels for Data-Driven Sense-Making and Decision Support

From Sensors to Sentinels to Sense

(for any application domain with streaming data from sensors)

- New knowledge and insights are acquired by monitoring and mining actionable data from all digital inputs (**Sensors!**)
- Alerts are triggered **autonomously, without intervention** (if permitted), applying machine learning and actionable business decision rules for pattern detection and diagnosis. (**Sentinels!** = embedded machine learning / data science algorithms, at the point of data collection = trained to minimize False Positives and "Alarm Fatigue")
- "Smart Sensors" (powered by Machine Learning-enabled sentinels) will therefore deliver actionable intelligence (**Sense!**)

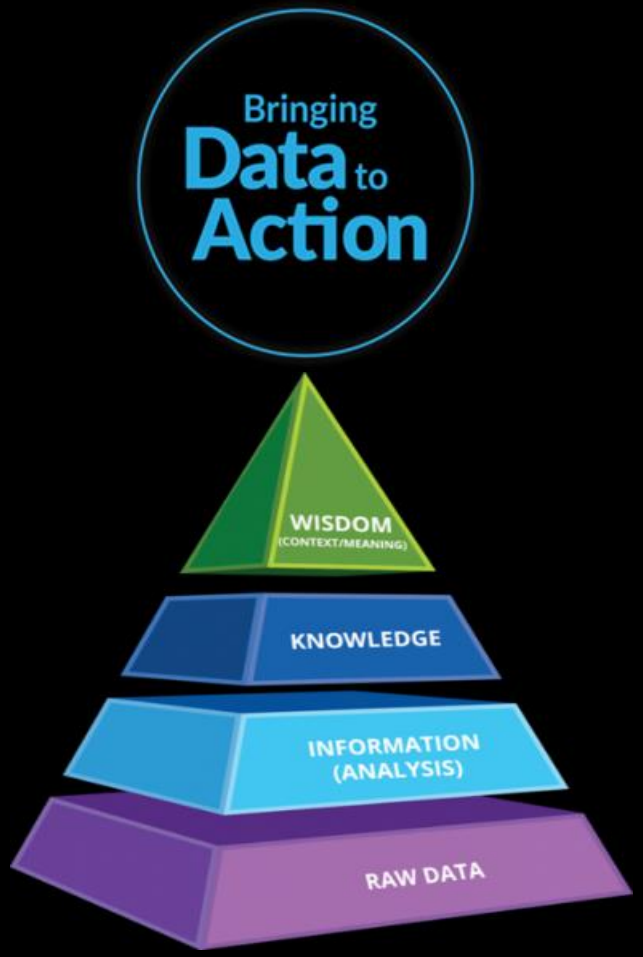
The Data Analytics Value Chain in 4 Steps

The Data Analytics Value Chain – go from Digital Data to Information to Knowledge to Insights (Decisions and Action)

- ✓ **From Sensors** (Measurement & Data Collection) …
  - ***… Big Data (Deep, Fast, Wide)***
- ✓ **to Sentinels** (Monitoring & Alerts = Information) …
  - ***… Machine Learning***
- ✓ **to Sense-making** (Knowledge & Insight Discovery) …
  - ***… Data Science***
- ✓ **to Cents-making** (Your Applications of Data = Action!)
  - ***… Analytics***
  - ***… Productizing / Monetizing your Data***

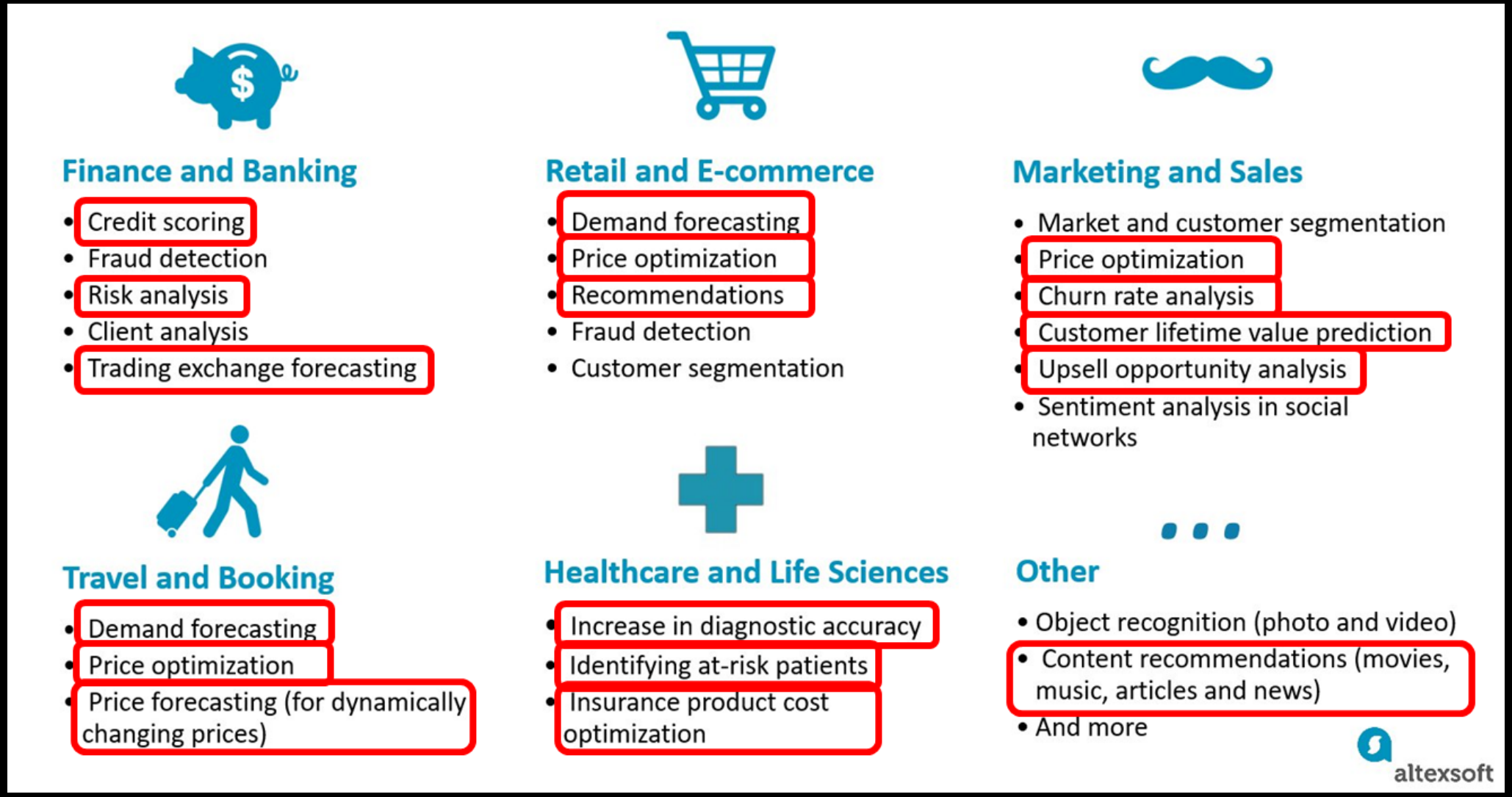
OUTLINE

- Preliminaries
- Having the “Talk”
- Self-Driving Enterprise
- **1001 Analytics Things to Do**



http://bizbench.com/wisdom-pyramid/

Predictive Analytics is **everywhere** in Business Data Science and Machine Learning Applications



Source for graphic: https://www.altexsoft.com/blog/datascience/machine-learning-strategy-7-steps/

“1001 Free Things” You Can Do with Your Data – 1

https://datamakespossible.westerndigital.com/over-1001-free-things-to-do-with-data/

General things to do with data:

Analytics product development	Machine intelligence
Anomaly/Outlier/Novelty discovery	Machine Learning (ML)
AI (= the ultimate data consumer!)	Mixed Reality (MR)
Association discovery	Natural Language Generation (NLG)
Augmented Reality (AR)	Natural Language Understanding (NLU)
Autonomous application systems	Network (Link) analysis
Bias detection in algorithms	Optimization
Blind source separation	Pattern discovery and recognition
Causal (Explanatory) factor analysis	Personalization
Classification	Predictive analytics (forecasting)
Computer Vision (CV)	Predictive maintenance
Content generation	Prescriptive analytics (optimization)
Correlation discovery	Prescriptive maintenance
Correlation discovery	Principal Component Analysis (PCA)
Data-driven decision support	Rare event detection
Deep Learning (DL)	Recommendations (recommender systems)
Diagnoses	Regression analysis
Document labeling	Reinforcement learning (gamification)
Edge analytics	Robotic Process Automation (RPA)
Embedded analytics	Segmentation
Fraud detection	Streaming analytics
Image classification	Supervised learning (classification, diagnosis)
Immersive Reality experiences	Text summarization
Independent Component Analysis (ICA)	Time series anomaly detection
Industry 4.0	Topic modeling
Intelligent edge	Unsupervised learning (pattern discovery)
Link discovery	Virtual Reality (VR)

Specific things to do with data:

4-D Printing	Intelligent data management
A/B testing	Intelligent search
Alert fatigue mitigation (fewer false positives)	Intent detection
Anti-money laundering	Inventory optimization and prediction
Application-specific dataset recommendation	IoT (Internet of Things) sensor fusion
Automated data labeling and tagging	IoT edge & streaming analytics
Automated feature importance ranking	Legal e-discovery
Automated image/video captioning	Legislation & Policy analytics
Automated question answering	Legislation loophole detection
Automated question generation	Literature-based discovery
Automatic data integration	Location intelligence
Automatic data quality assessment	Machine translation
Autonomous disaster response planning	Map routing
Call center automation	Market discovery
Change-point detection in streaming data	Market trends detection and prediction

Chronic illness prediction	Medical diagnosis
Compliance verification	Missing alerts recovery (false negatives)
Connected products/vehicles	Missing value imputation
Content labeling	Mortgage approval acceleration
Conversational AI (chatbots)	Mortgage risk analysis
Cross-sell/Up-sell opportunities discovery	Multichannel customer analytics
Customer behavior analytics	Music classification
Customer experience optimization	Music generation
Customer journey modeling and analytics	New materials discovery
Customer service automated text response	Object detection and classification
Customer/Employee churn prediction	Omnichannel customer engagement
Customization of portals and experiences	Omnichannel marketing attribution
Cyber behavior analytics	Operational efficiencies discovery
Cyber threat detection	Pandemic spread prediction
Data (metadata) enrichment	Personalized training recommendations
Data loss detection	Precursor event analytics
Data Product Development	Predictive crime trends
Data security verification	Predictive customer choice selections
Data usage tracking	Predictive pricing
Data use case recommendation	Predictive product innovation
Dataset recommendation	Procurement/invoice/contract auditing
Digital Twins	Product demand prediction
Drug discovery	Risk analysis (detection and prediction)
Drug interaction discovery	Self-driving vehicles
Email filtering, ranking, sorting automation	Sentiment analysis in text
Email phishing detection	Shipping optimization and prediction
Email spam detection	Smart cities
Entity cross-identification	Smart energy grid
Entity disambiguation	Smart farms
Facial recognition	Smart homes
Fake content detection (news, images, videos)	Smart manufacturing
Food security tracking	Smart transportation
Healthcare patient behavior analytics	Speech (Speaker) recognition
Hierarchical segmentation	Sports analytics
Home price prediction	Supply chain prediction & optimization
Human Resources (HR, People) analytics	Sustainability analytics
Humanoid robots	Time series change-point detection
Hyper-dimensional data indexing	Traffic congestion mitigation
Hyper-personalization	User experience optimization
Illicit trafficking detection	Virtual digital assistants
Image generation from captions	Voice analytics (content, sentiment)
In-chip sensor analytics	Voice Assistants
	Voice-based search

“1001 Free Things” You Can Do with Your Data – 2

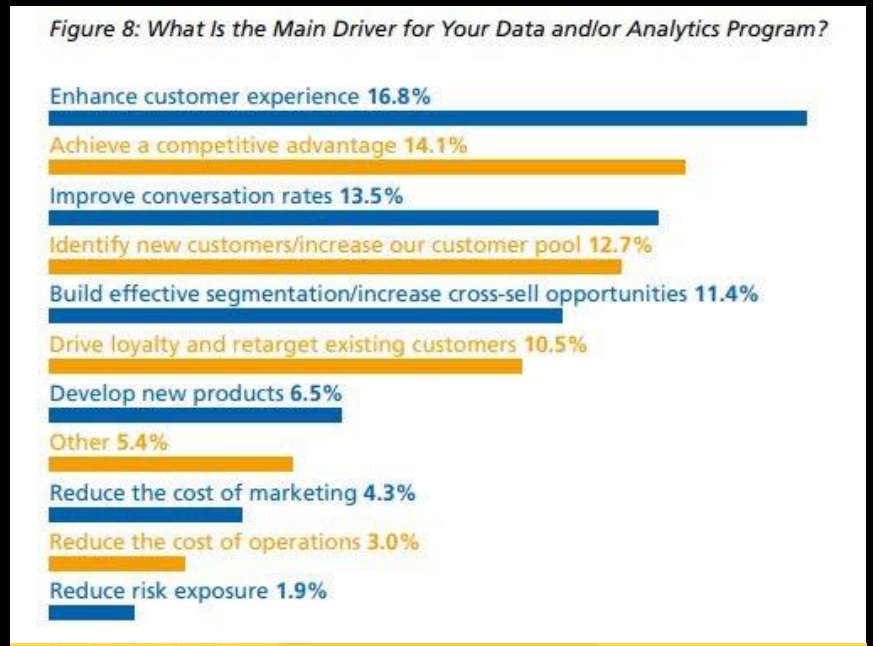
https://datamakespossible.westerndigital.com/over-1001-free-things-to-do-with-data/

Example – 15 use cases of predictive analytics on streaming business data:

1. Real-time credit risk prediction.
2. Real-time fraud risk prediction.
3. Real-time personalized customer interactions.
4. Real-time context-specific and location-sensitive product and/or content marketing to consumers.
5. “Do Not Pay” classification on fraudulent insurance claims prior to payment.
6. Real-time determination of benefits eligibility to mitigate underwriting fraud.
7. Detection of insurance rate evasion tactics within the policy quote process.
8. Optimal actuarial price determination at the point of policy-quote decision-making.
9. Health risk prediction at the point of healthcare decision-making.
10. Real-time detection of anomalous and adversarial cyber network behaviors.
11. Stop data breaches before they happen.
12. Stop illegal funds transfers before they happen.
13. Stop non-compliant business transactions before they happen.
14. Optimize supply chain and warehouse product flows: position the right products in the right quantities at the right locations just-in-time.
15. Predict product demand and pricing by finer levels of product subcategories.

Data Science, Analytics, and AI Use Cases in Various Industries & Sectors

- Agriculture: https://bit.ly/2UG7OYl
- Administration: https://bit.ly/2YbUAYa
- Banking: http://bit.ly/2LwvOZ9
- Construction: http://bit.ly/2X3hMnu
- Customer Support: http://bit.ly/2WbvVAm
- Design: http://bit.ly/2pN0znt
- Energy and Utilities: https://bit.ly/3de1aSi
- Finance: https://bit.ly/3hGvLLs
- Gaming: https://bit.ly/2YSz4GO
- Government: https://bit.ly/3egIIJM
- Healthcare: https://bit.ly/2UYD47z
- Human Resources: https://bit.ly/37FQdYg
- Insurance: http://bit.ly/2NPRim4
- Manufacturing: https://bit.ly/3fCpO0h
- Marketing: https://bit.ly/2Ybj8Au
- Pharma: http://bit.ly/35KgP8R
- Production: https://bit.ly/2YNgy2F
- Retail: http://bit.ly/2uVwodm
- Sales: https://bit.ly/2BhjGf8
- Telecom: https://bit.ly/3eatLci
- Transportation: https://bit.ly/30UI6Us
- Travel: https://bit.ly/3hGuvYT
- and more… : https://bit.ly/2LJKE21



10 common uses for Machine Learning applications in business

1. Real-time Chatbot Agents (NLP, NLU, NLG, Conversational AI)
   - Customer Service, Call Centers, Financial Services, Voice Assistants (Siri, Alexa, Cortana, Google Assistant)
2. Decision Support (Actionable Insights, Prescriptive Analytics, B.I., Dashboards)
   - Healthcare, Business, Agriculture, Manufacturing, Maintenance, Climate Action, Security
3. Customer Recommendation Engines (Predictive Analytics, Personalization, Sentiment Analysis)
   - Content-based, Context-based, Collaborative Filtering, Shopping Preferences, E-Commerce
4. Customer Churn Modeling (Predictive Behavioral Analytics, Pattern Discovery)
   - Mobile / Telecom Companies, Music / Movie Streaming Services, Media Companies, Financial / Banking
5. Dynamic Pricing (Context-based, Dynamic Data-Driven, Supply-Demand Monitoring)
   - Surge / Demand Pricing, Airline Tickets, Rideshare / Taxis, Retail, Transportation, Energy, CPG
6. Market Research and Customer Segmentation (Predictive Inventory Planning, Predictive Analytics 2.0)
   - Retail, Inventory Placement, Durable Goods Manufacturing, Buying Patterns, Emerging Trends
7. Fraud Detection (Graph / Linked Data Analytics, Anomaly Detection, Streaming Real-time Analytics)
   - Financial Services, Travel, E-Commerce, Retail, Gaming, Claims (Benefits & Insurance), Phishing / SPAM
8. Image Understanding (Recognition & Classification, Computer Vision, Deep Learning, Image Captioning)
   - Robotics, Retail (Shopping Carts, Shelf Inventory), Autonomous Vehicles, Safety / Risk Management
9. Operational Efficiency (Process Discovery, Robotic Process Automation RPA, Digital Process Automation)
   - Finance, Facilities & Equipment Maintenance, I.T. Operations, Documents/Forms Processing, DevOps
10. Information Retrieval & Extraction (NLP, Document Understanding, Text Analytics, RPA Document Processing)
    - Finance, Legal, Contracts, Customer Service, Product Information Management (PIM)

https://searchenterpriseai.techtarget.com/feature/10-common-uses-for-machine-learning-applications-in-business

AI (Machine Learning, NLP) Document Processing is a subtle but powerful use case for Business Analytics

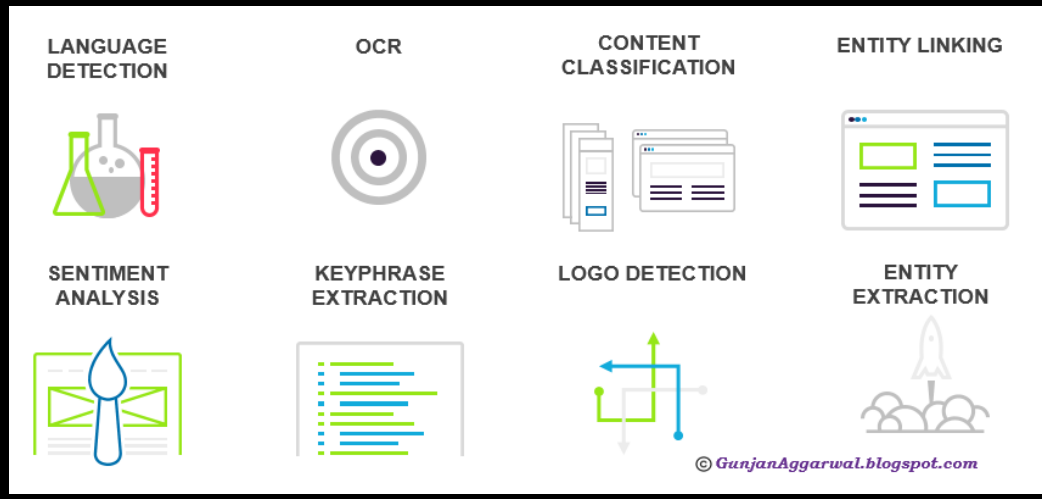
https://searchenterpriseai.techtarget.com/feature/AI-document-processing-remains-a-subtle-but-powerful-use-case

AI-enhanced Document Processing appears in at least 5 items on the preceding slide:

2. Decision Support
6. Market Research and Customer Segmentation
7. Fraud Detection
9. Operational Efficiency
10. Information Retrieval & Extraction

Use cases (Marketing, Medical, Legal, Contracts, Claims):

- Content Summarization (Document Understanding)
- Document Categorization (sorting)
- Document Classification (labeling)
- Entity Discovery, Extraction, & Analysis (Links, Associations)
- Insights Discovery for Decision Support
- Key-phrase (Information) Retrieval & Extraction
- Link Analysis (Graph-based e-Discovery)
- Logo Detection & Extraction
- Sentiment Analysis
- Trend Analysis

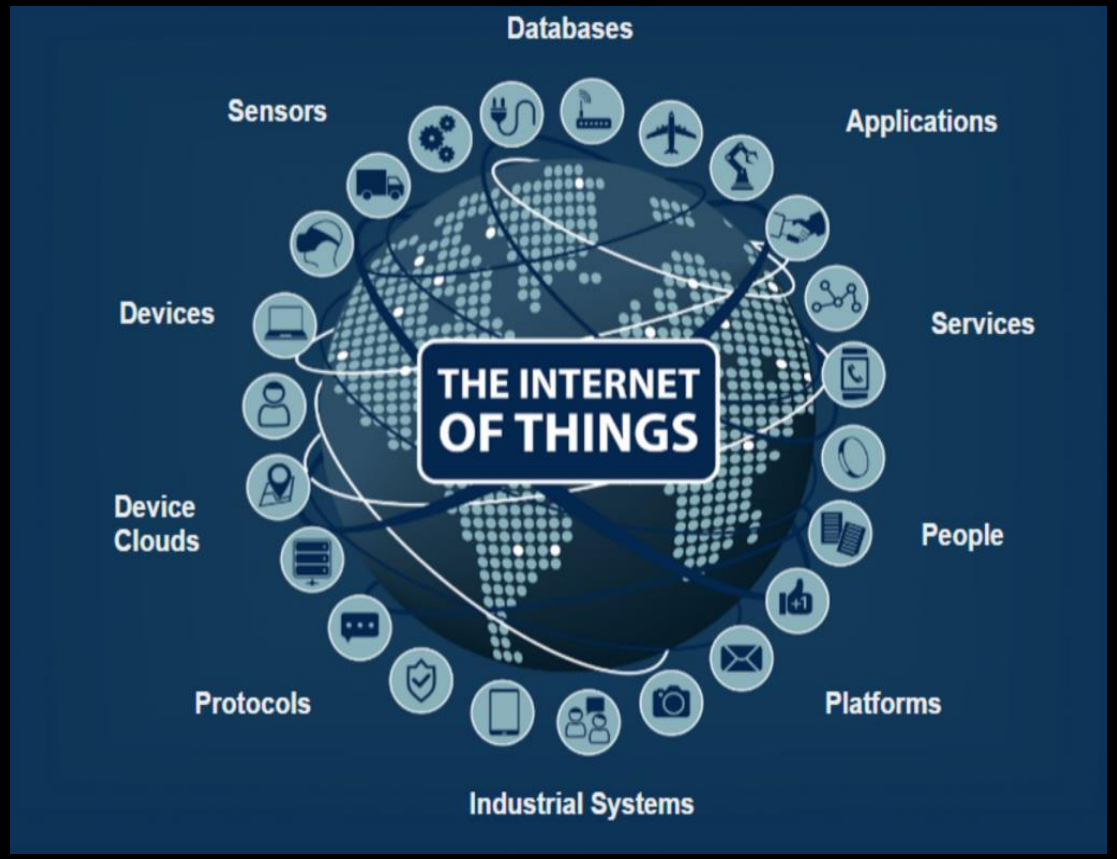


http://gunjanaggarwal.blogspot.com/2018/06/components-of-ai-based-document.html

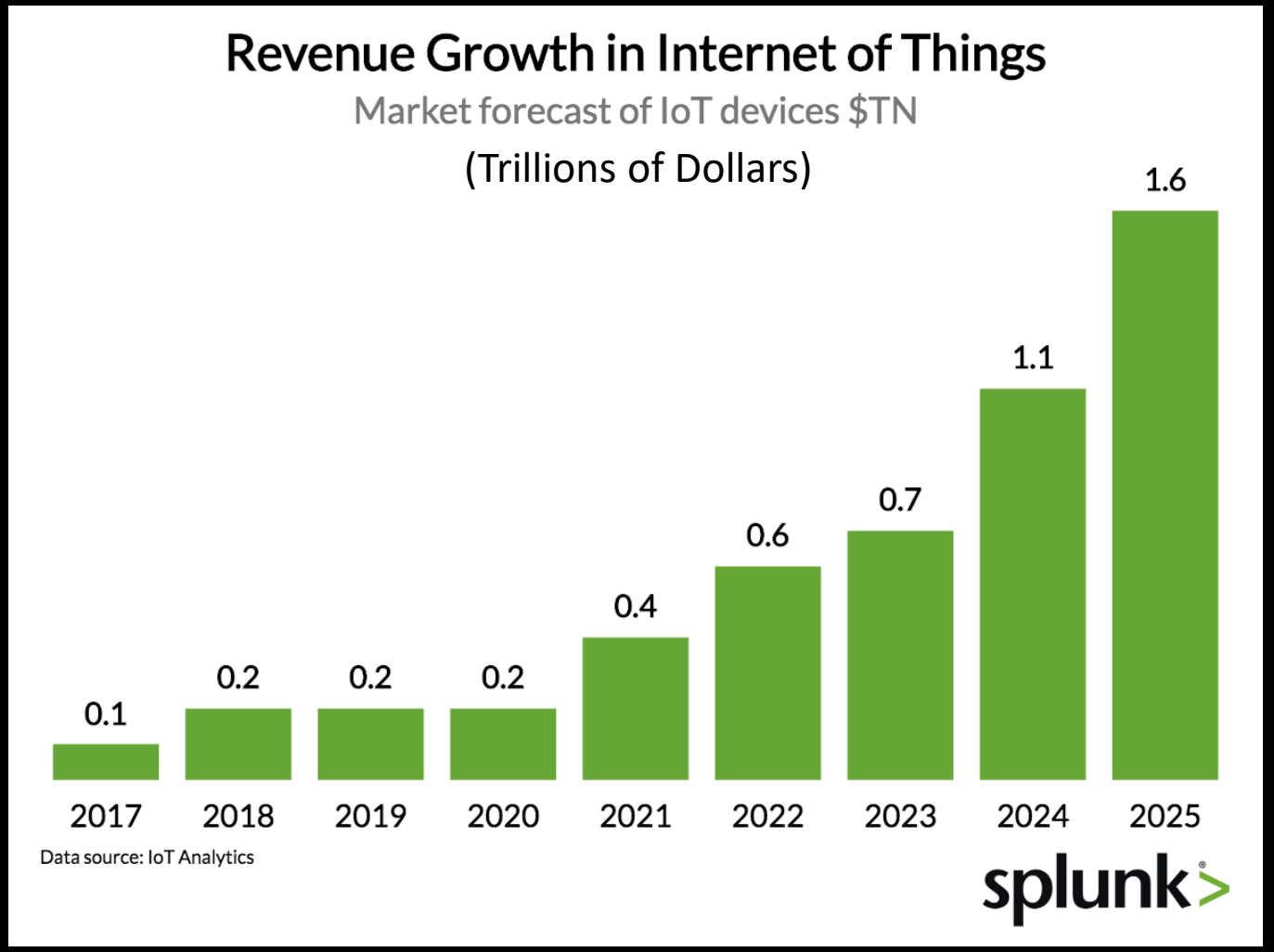
While we are on the subject of massive data and value-packed things you can do with all that data…
Don’t forget the Internet of Things (IoT)!
…including the Industrial IoT (IIoT)

I assure you that I have not forgotten the IoT!

Incoming… Explosive Growth in IoT and IIoT!

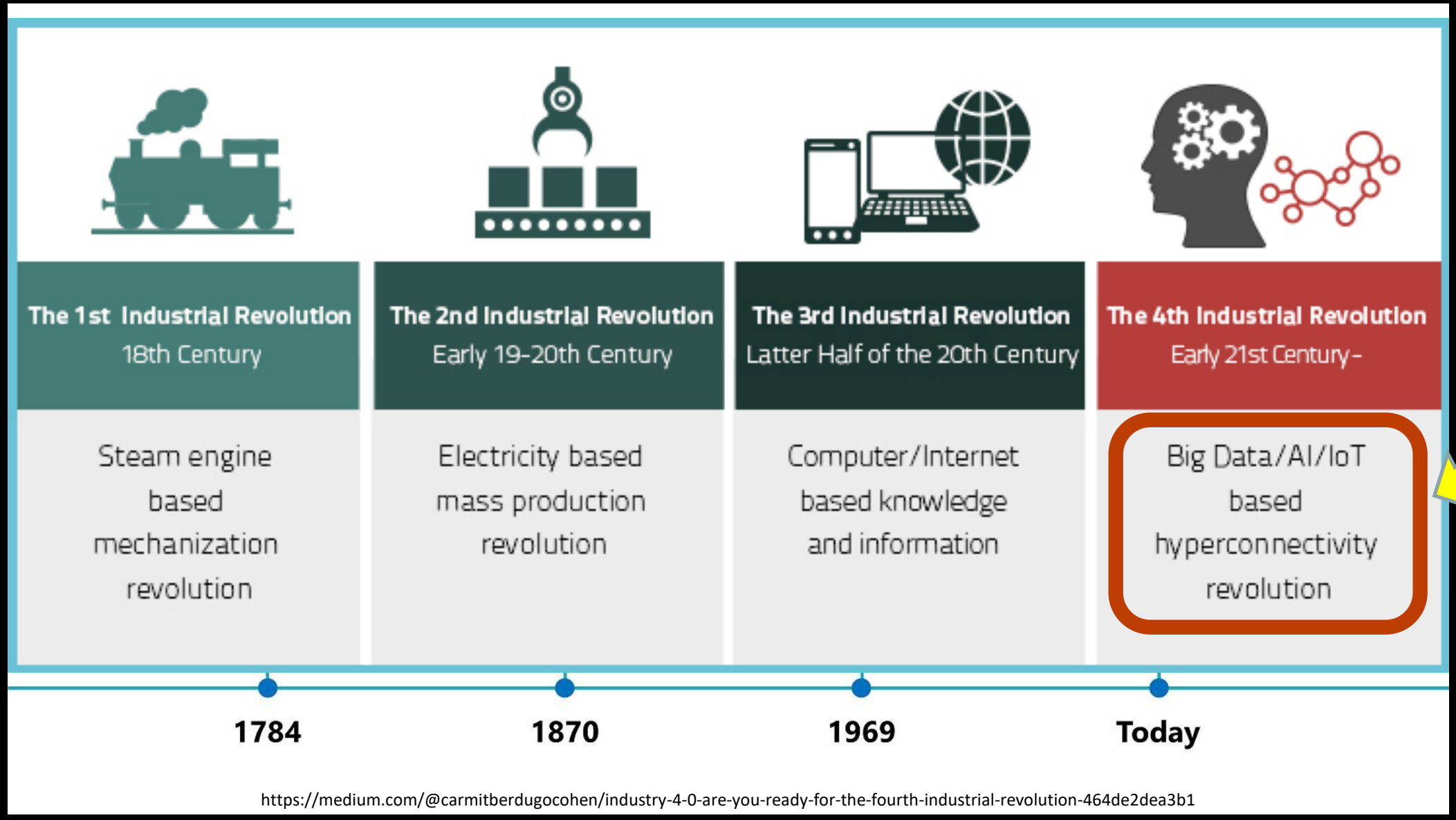


https://www.nexigen.com/blog/internet-of-things-smart-home-explained/



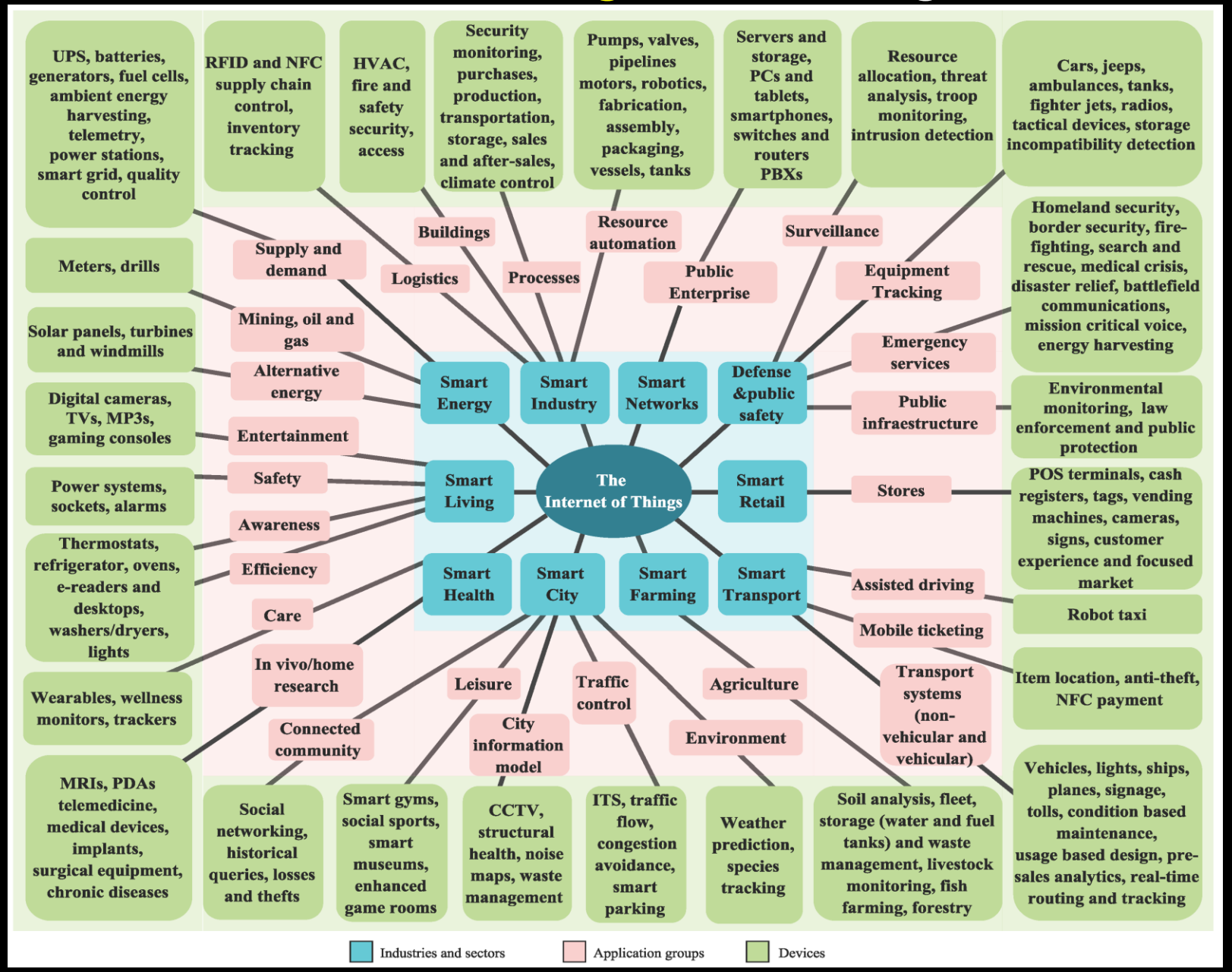
https://priceonomics.com/the-iot-data-explosion-how-big-is-the-iot-data/

The 4th Industrial Revolution = Industry 4.0
or The 4th Innovation Revolution = Innovation 4.0



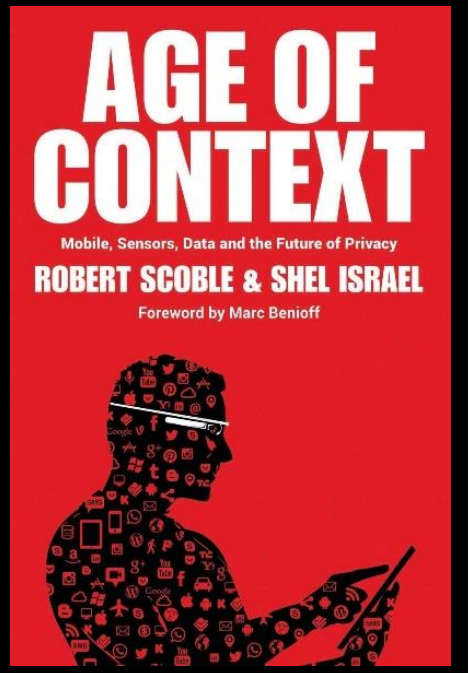
https://medium.com/@carmitberdugocohen/industry-4-0-are-you-ready-for-the-fourth-industrial-revolution-464de2dea3b1

The Internet used to be a thing. Now, things are the Internet.

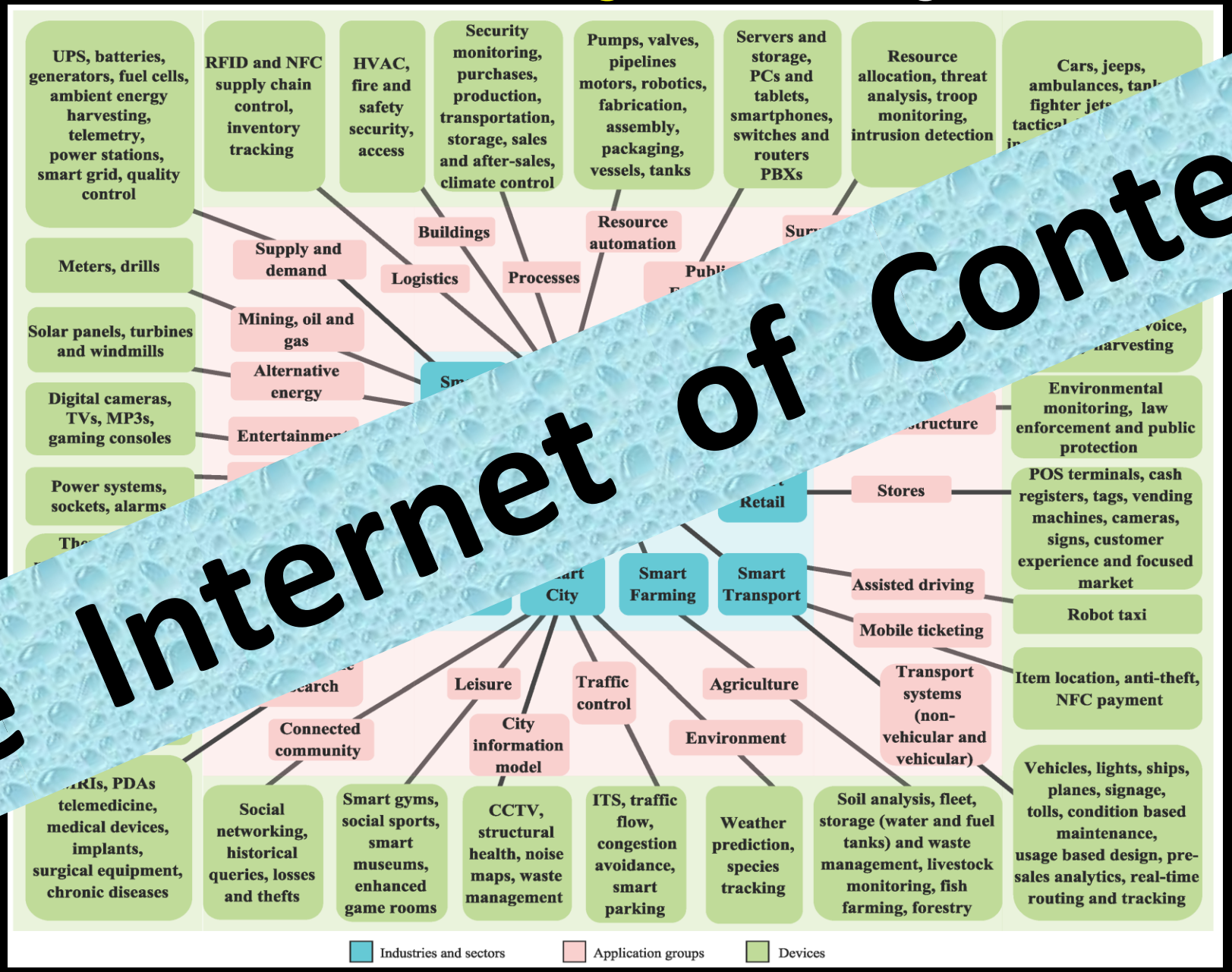


Source for graphic: http://www.mdpi.com/1424-8220/16/10/1644

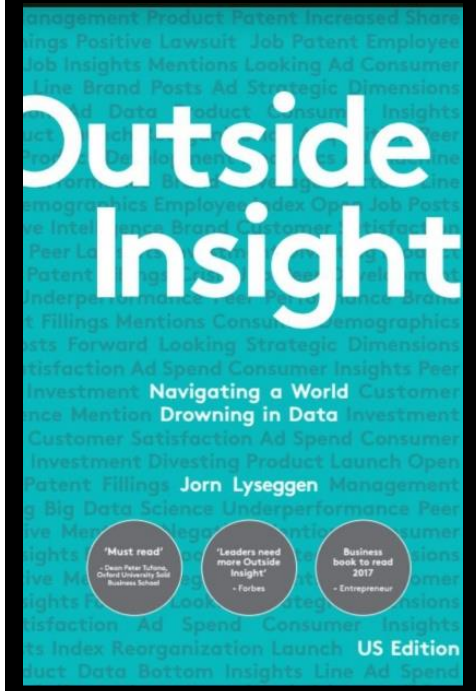
The Internet used to be a thing. Now, things are the Internet.



https://amzn.to/34utTz7



The Internet of Context



https://amzn.to/317XV9L

Source for graphic: http://www.mdpi.com/1424-8220/16/10/1644

The Internet used to be a thing. Now, things are the Internet.

IoT is more than just monitoring with sensors and data!

- Deploying IoT sensors everywhere guarantees a data deluge. **Ouch!**
- You don't need more information. What you need is more insights. **Truth!**
- Consequently, **Smart IoT deployments** are essential.
- How IoT deployments are "Smart" = Cognitive, Contextual, Predictive, and…
  - *they are informed by deliberate decisions about where and when more data (and more insights) are needed for your organization;*
  - *they are strategic, to meet desired outcomes and mission objectives.*
- Therefore, you need a Strategy – specifically, an **Observability Strategy** – to deliver mission-critical **Insights-as-a-Service**.
- Monitoring is WHAT you do, and Sensors are HOW you do it.
- **However… Observability is WHY you do it.**
- **Insights from data are the REWARDS for doing it.**
- **Add some APIs to sensors – You will gain Context and Insights-as-a-Service from your Observability Strategy.**

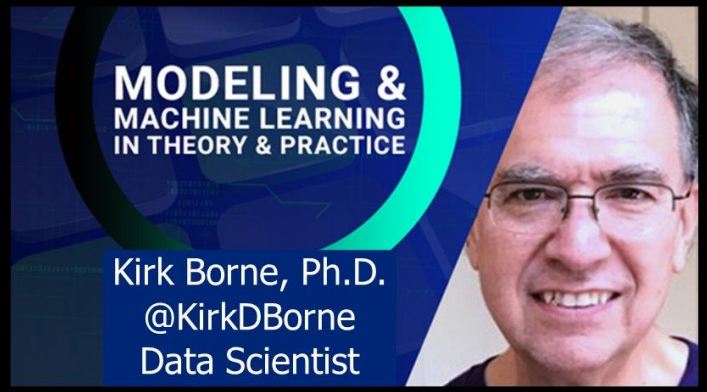
Insights Discovery

Things to Do with Insights-as-a-Service

- Cognitive Analytics = Discover the right questions in your data – That's cognitive!
  - *"What question should I be asking when I see something funny in my data?"*

  http://bit.ly/2MLZEj8
- Precursor Analytics = Detect early-warning signs of events – That's Forecasting-as-a-Service.
  - *"What are these changing & emerging trends in the data telling me?"*

  https://twitter.com/KirkDBorne/status/1359495081531756548
- Sentinel Analytics = Use sensors and data to "watch" the things that should be watched, to be assured that things are working fine – That's Compliance & Comfort.
  - *"I'm picking up good vibrations."*



Course Format:
1) Two sessions
2) Four content modules per session

Session 2: Typical and Novel (Atypical) Applications of Machine Learning (ML) Models

- A Fishy Example of Cost-Sensitive Classification
- Exploratory Data Analysis: Some Simply Successful Insights-Driven Models
- A 12-step Analytics Program in Healthcare and Medicine
- ML and AI Modeling Opportunities in Business and Marketing
- Data Literacy Exercises: Strengthening Your Data Science Abilities
- Surprise Discovery in Regression Analysis
- Neural Networks in Climate Modeling
- ICA vs. PCA: The Cocktail Party Problem
- Graph Mining: Connecting the Dots that Aren't Connected
- Forecasting 2.0: Beyond Traditional Forecasting
- Clustering Analysis: Down to Earth, and Up to Space
- Association Mining for Predictive Modeling
- The Ways of Bayes: Classification, Markov Models, Missing Value Imputation, Causal Analysis
- Precursor Analytics with Statistical Clustering
- The Internet of Context: Forecasting-as-a-Service
- Matching ML Algorithms to Business Problems: Steps to Data Analytics Mastery
- Course Wrap-up: The Keys to a Successful Data Science Career

@KirkDBorne

Your Stellar Analytics Scorecard for Data Innovation and Analytics Mastery in the Self-Driving Enterprise

1) **Streaming Data Analytics:** Real-time access to, interaction with, and discovery from data…
   - Detects POI (Person, Pattern, Product, Process, or Point Of Interest)
   - Detects BOI (Behavior Of Interest from any "dynamic actor")

Detect Emerging Classes (Segments), Correlations, Surprises (Anomalies), Links (Associations) in incoming data.

2) **Team Analytics:** Culture of experimentation, diversification, collaboration, data-sharing, data reuse, and data democratization

3) **Edge Analytics:** Locality in Time, at the moment of data collection (What else is happening now?)

4) **Location Analytics:** Locality in Geospace, in that context (What else is happening at that place?)

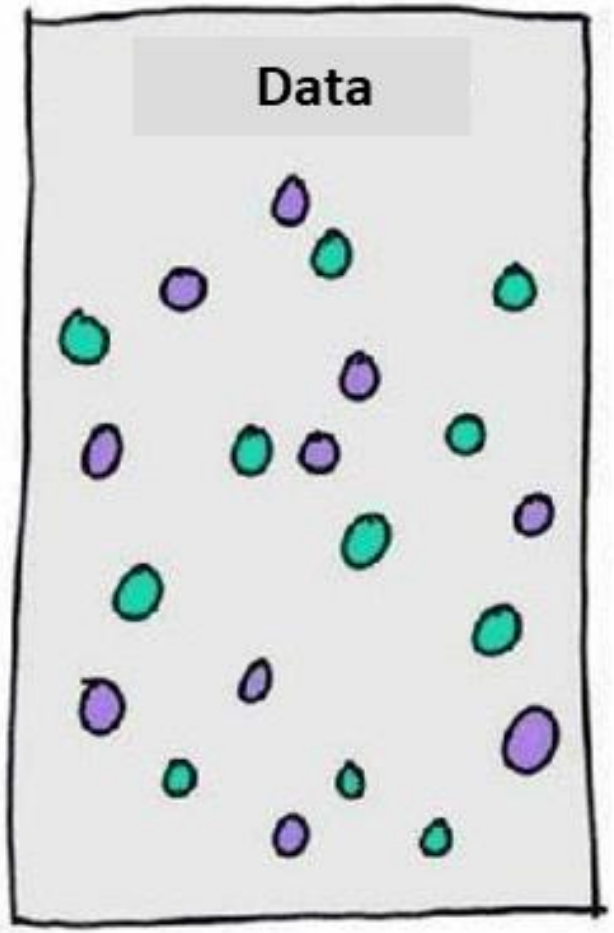
7) **Learning Business System Analytics:** Data-driven knowledge-generation business processes, with continuous feedback, learning, and improvement – embedded in daily business practice (REF)

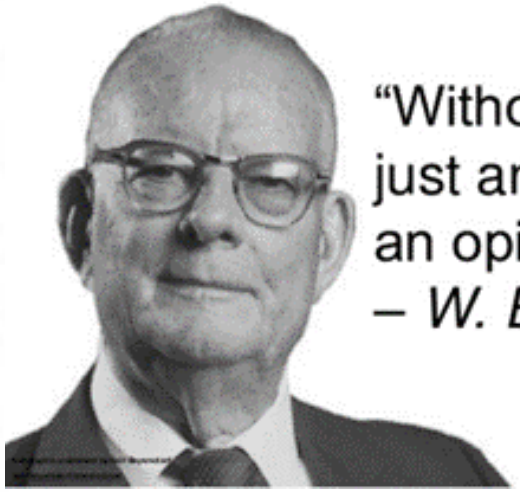
6) **Agile Analytics:** Analytics By Design, Iterative, Build Proofs of Value, Fail-fast to Learn Fast, MVP (and MLP = minimum lovable product), CI / CD (Continuous Integration / Delivery)

5) **Related-entity Analytics:** Locality in Data Feature Space (What else is like this entity / event?)

(REF: https://en.wikipedia.org/wiki/Learning_health_systems)

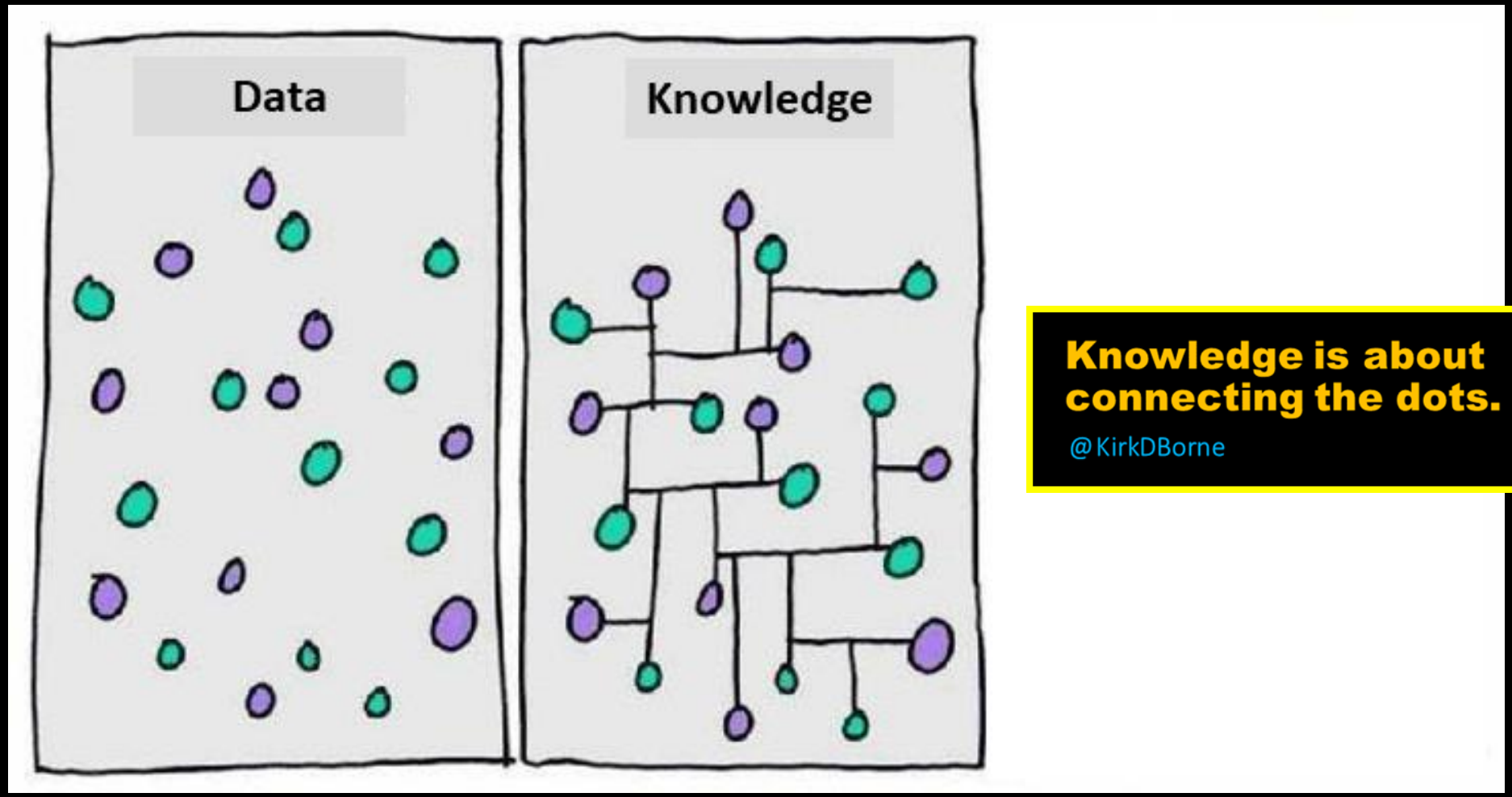
Data



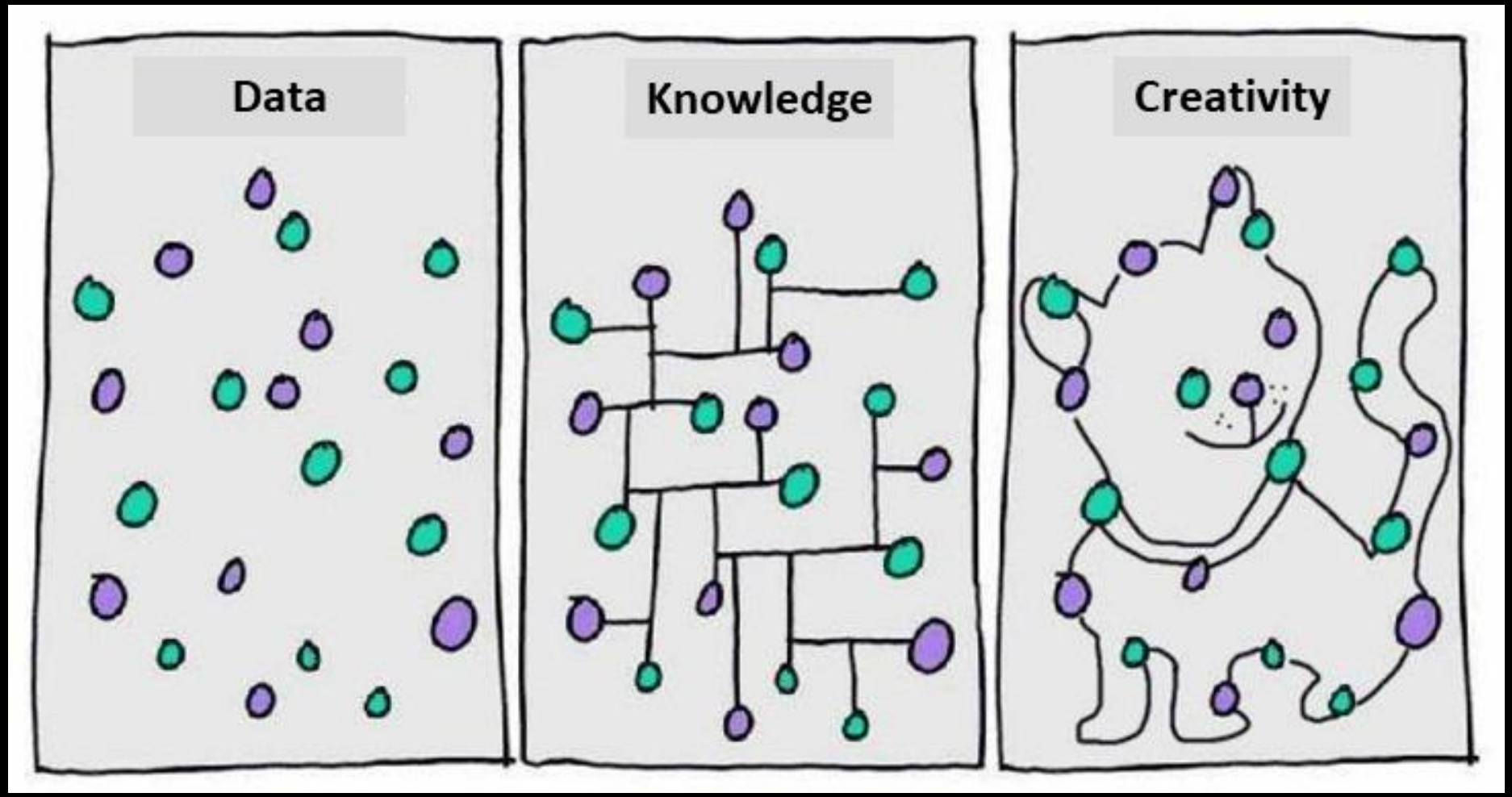


"Without data you're just another person with an opinion."
– *W. Edwards Deming*

Data **Data Science (KDD)**



Data Data Science **Create & Do Something**



Bringing Data to Action

http://bizbench.com/wisdom-pyramid/

Thank you!

Kirk Borne

@KirkDBorne

These slides here: http://www.kirkborne.net/datadotworld2021/