



LOGO

LẬP TRÌNH CHO KHOA HỌC DỮ LIỆU

Bài 1. Tổng quan về khoa học dữ liệu

Nội dung

1

Khoa học dữ liệu là gì

2

Ứng dụng của khoa học dữ liệu

3

Data scientist (nhà khoa học dữ liệu)

4

Các yếu tố phát triển khoa học dữ liệu

Khoa học dữ liệu là gì

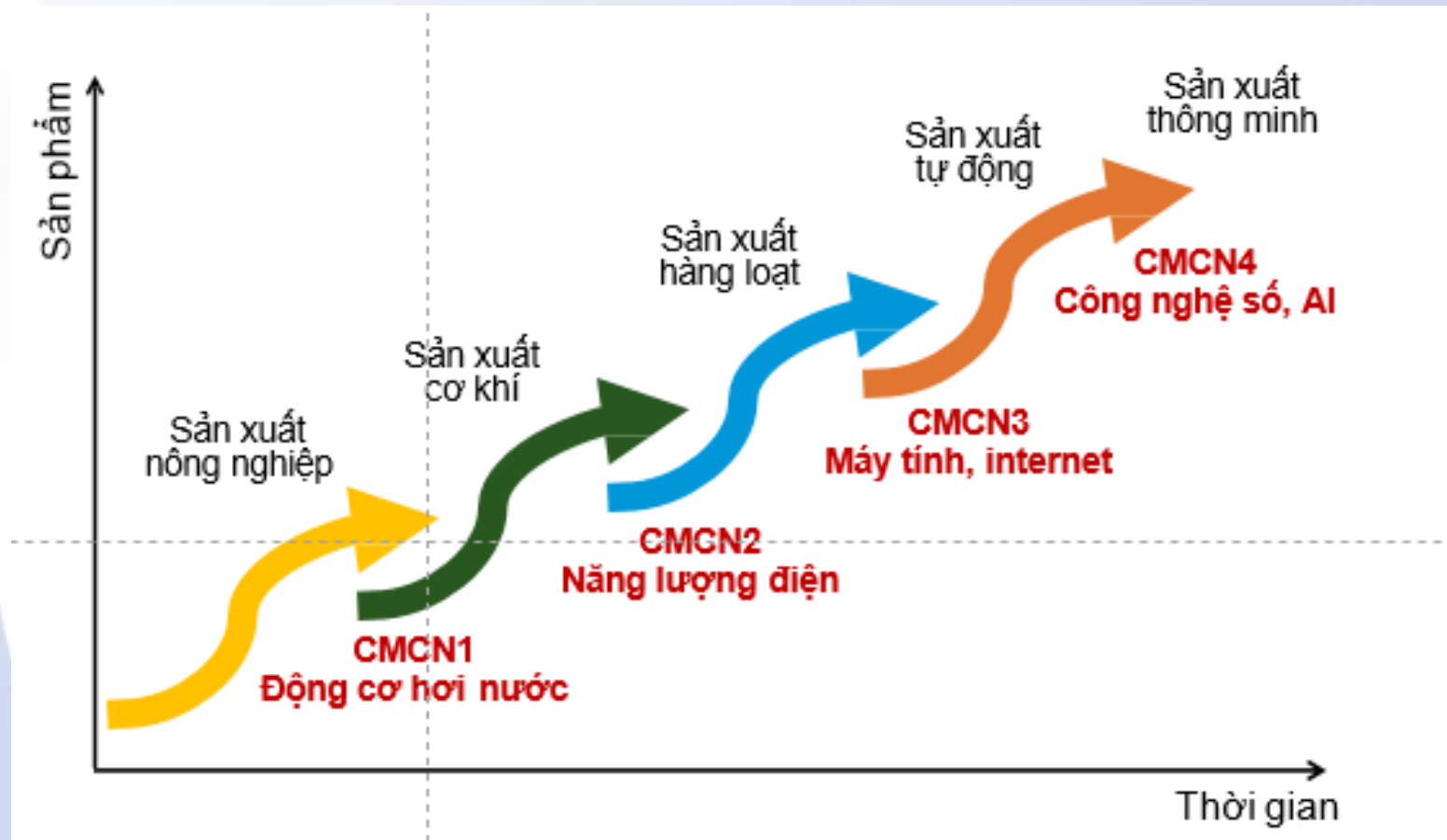
- Hầu hết các ngành khoa học từ xưa đến nay đều giải quyết vấn đề dựa trên **lập luận** và **tri thức**
 - Ngành toán: dựa trên các mệnh đề, công thức, lập luận... để chứng minh bài toán
 - Ngành vật lý: dựa trên các quan sát, thực nghiệm, tính toán,... kiểm chứng các giả thiết
 - Ngành hóa học:...
 - ...
 - Tagọi các ngành khoa học này là “knowledge-driven” (dẫn dắt bởi tri thức)
- Có ngành có chút ngoại lệ, ví dụ: **ngành xác suất**

Khoa học dữ liệu là gì

- Với quan điểm như vậy, tất cả những quan sát mà không được chứng minh chặt chẽ thường được cho là “không khoa học”
 - Chẳng hạn: chuồn chuồn bay thấp thì mưa
- Khoa học dữ liệu \neq Khoa học thông thường ở quan điểm: **tìm tri thức từ dữ liệu** (dẫn dắt bởi dữ liệu – “data-driven”)
 - Chúng ta rút ra tri thức bằng việc tìm tòi từ dữ liệu chứ không nhất thiết phải chứng minh nó
 - Tất nhiên tri thức tìm ra phải có tính ổn định (luôn có cùng kết quả nếu sử dụng cùng một phương pháp)

Khoa học dữ liệu là gì

Cách mạng công nghiệp lần thứ tư



Khoa học dữ liệu là gì

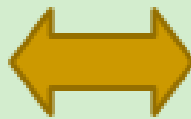
Cách mạng số hoá và physical-cyber systems

- **‘Phiên bản số’** các thực thể: Biểu diễn các thực thể bằng ‘0’ và ‘1’ trên máy tính (digital version)

Thí dụ: ô-tô, bệnh án điện tử...

- **Hệ thống không gian số-thể giới thực thể (cyber-physical system):** kết nối các thực thể và ‘phiên bản số’ của chúng.

Hành động trong
thế giới các thực thể

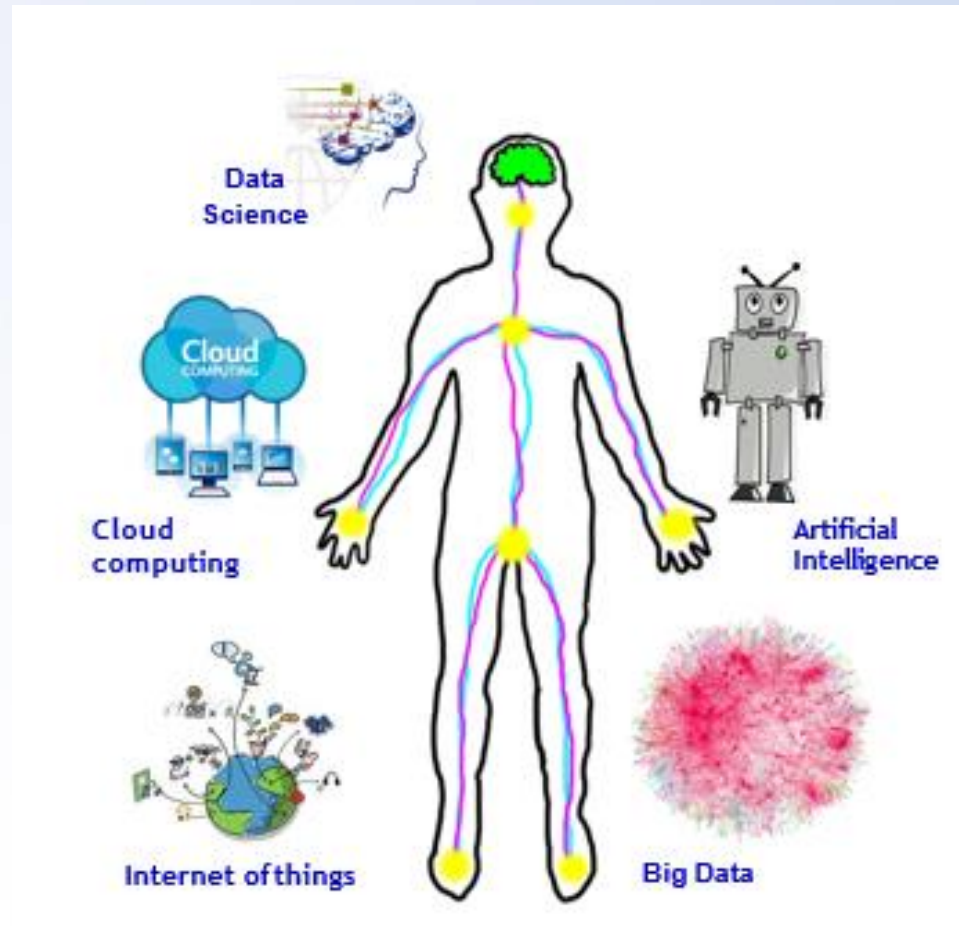


Tính toán, điều khiển
trên không gian số

Thay đổi phương thức sản xuất
Ảnh hưởng mọi lĩnh vực của xã hội

Khoa học dữ liệu là gì

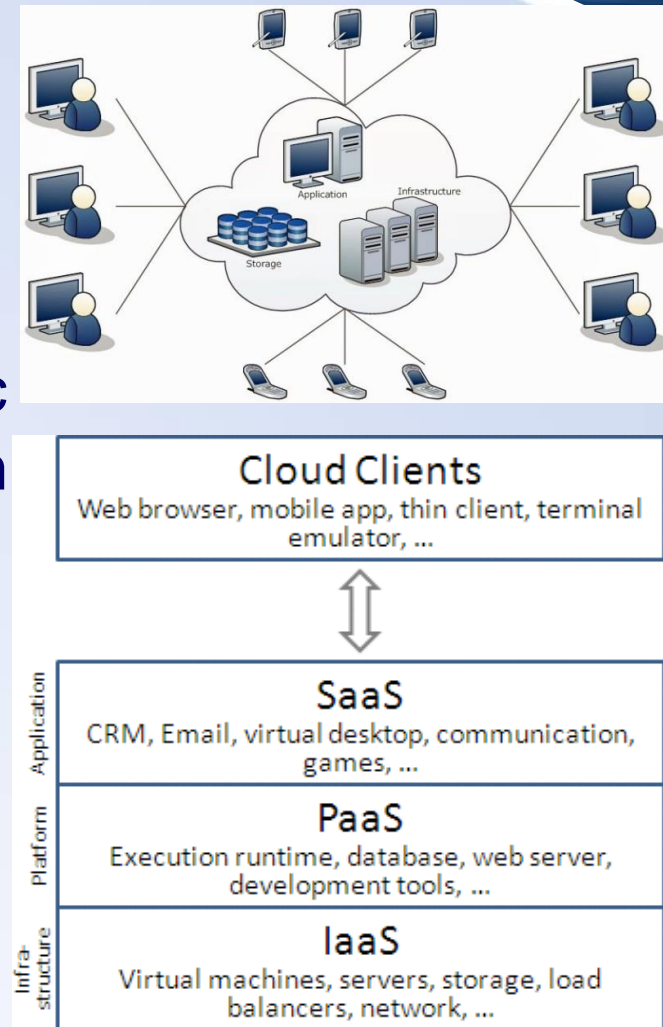
Đột phá trong chuyển đổi kỹ thuật số



Khoa học dữ liệu là gì

Cloud computing

- Điện toán đám mây: Lưu giữ và truy nhập dữ liệu và chương trình trên clouds qua Internet thay vì trên máy tính của người dùng.
- Dữ liệu được lưu giữ thường trực tại các máy chủ trên Internet và chỉ lưu trữ tạm thời ở máy khách.
- Cloud computing vs local computing
- Ví dụ:
 - Google drive, Google gmail
 - Apple iClouds
 - Dropbox

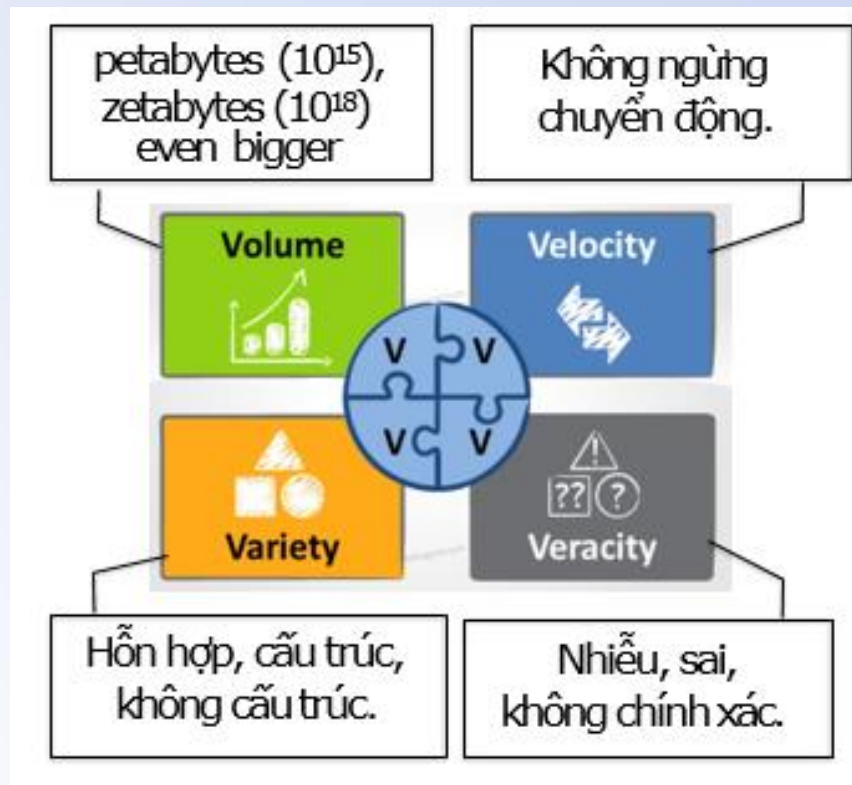
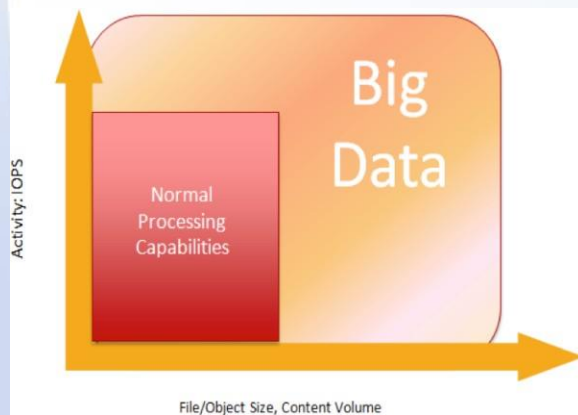


Internet of things (IoT)

Khoa học dữ liệu là gì

Big Data

Dữ liệu lớn nói về các tập **dữ liệu rất lớn** và/hoặc **rất phức tạp**, vượt quá khả năng xử lý của các kỹ thuật IT truyền thống.



Khoa học dữ liệu là gì

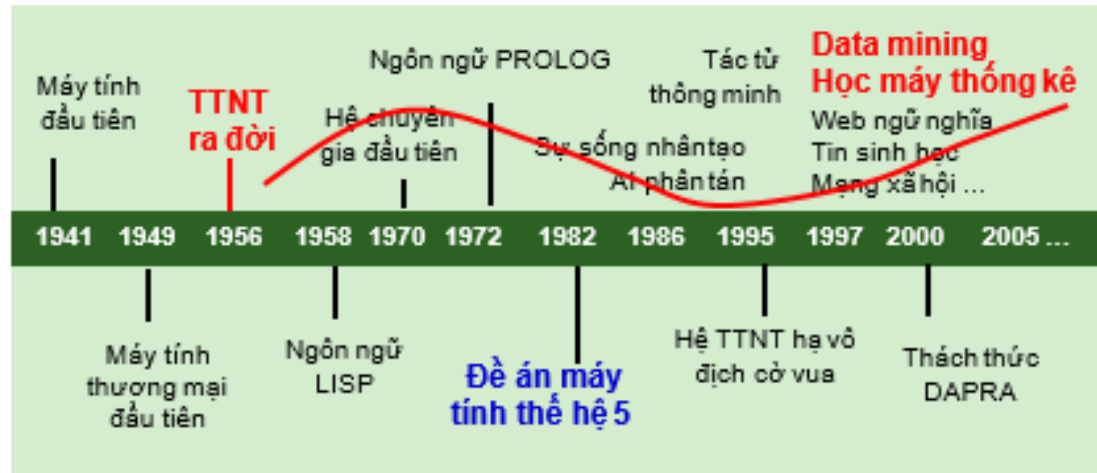
Artificial Intelligence – Trí tuệ nhân tạo



- Lĩnh vực làm cho máy (tính) hoạt động như có trí thông minh của con người (lập luận, hiểu ngôn ngữ, học tập...).
- Phép thử Turing là một cách để trả lời 'máy tính có biết nghĩ không?'

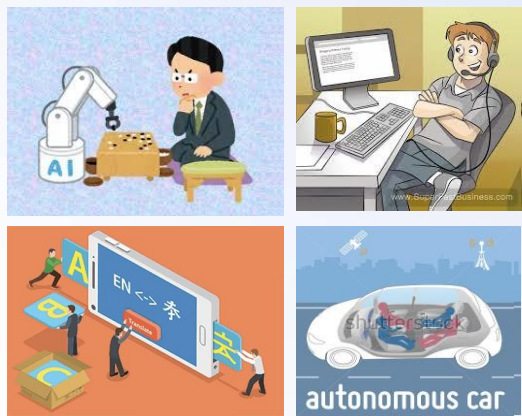


1912-1954

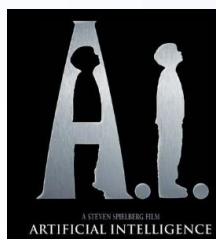


Khoa học dữ liệu là gì

Artificial Intelligence – Trí tuệ nhân tạo



- Lĩnh vực làm cho máy (tính) hoạt động như có trí thông minh của con người (lập luận, hiểu ngôn ngữ, học tập...).
- AlphaGo, hiểu ngôn ngữ, tiếng nói, chẩn đoán ung thư, ô-tô tự lái...



=



+



+



Hầu hết đột phá gần đây của AI dựa vào học máy (machine learning).

Khoa học dữ liệu là gì

Vài định nghĩa về Khoa học dữ liệu?

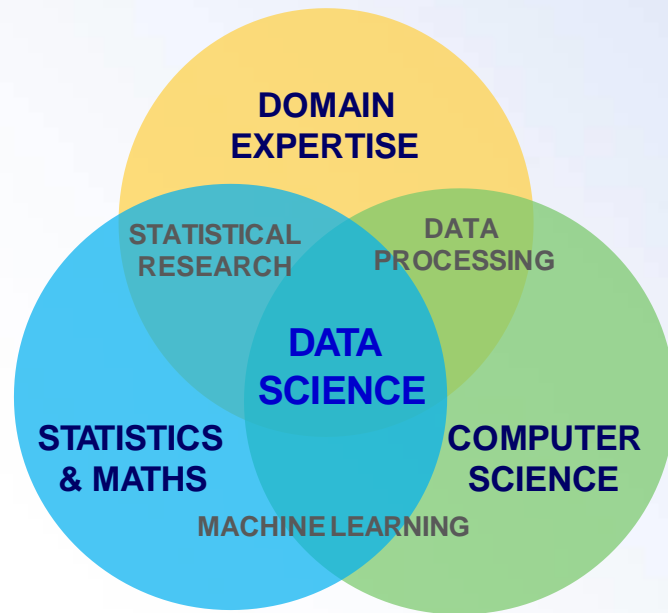
- Chưa có sự thống nhất.

NIST (National Institute of Standards and Technology)	Data science is extraction of actionable knowledge directly from data through a process of discovery, hypothesis, and hypothesis testing Trực tiếp trích rút tri thức hành động từ dữ liệu qua quá trình phát hiện, thiết lập và kiểm nghiệm các giả thiết.
Microsoft	Data science is about using data to make decisions that drive actions. Dùng dữ liệu tạo quyết định dẫn dắt hành động

Thay đổi: data analysis → data analytics → data science

Khoa học dữ liệu là gì

Data science

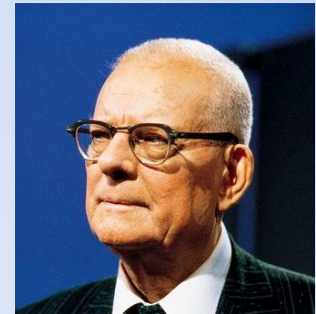


Source: Palmer, Shelly, Data Science for the C-Suite
New York: Digital Living Press, 2015

“In God we trust.
All others bring
data”.

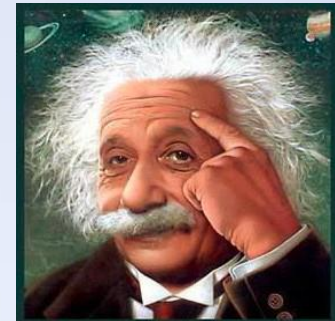
“Ta tin Thượng đế.
Ngoài ra, là dữ
liệu”.

W.E.D eming

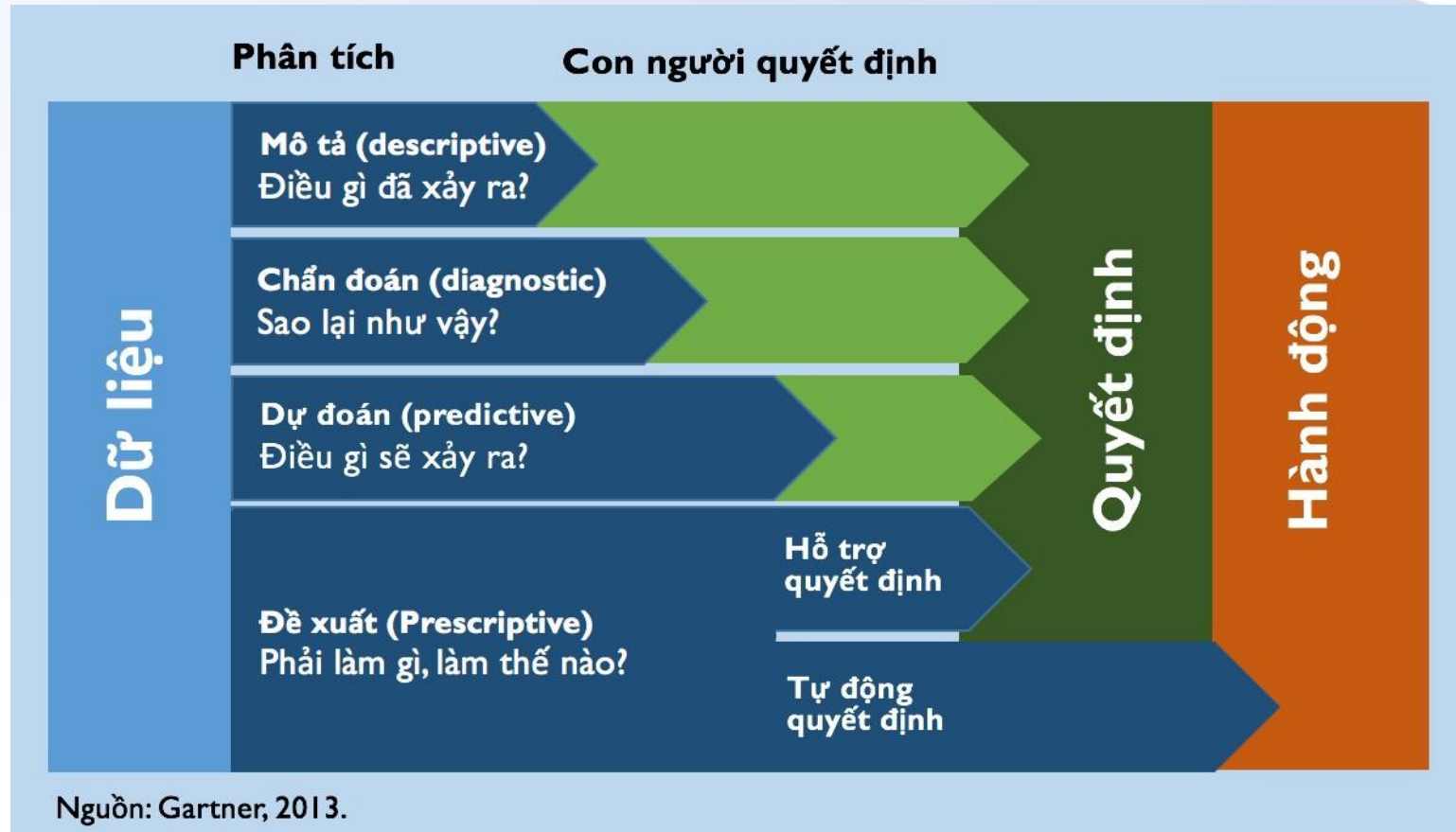


“We cannot solve problems
by using the same kind of
thinking we used when we
created them”

Ta không thể giải quyết các
vấn đề với chính cách nghĩ ta
đã dùng khi đặt vấn đề
Albert Einstein

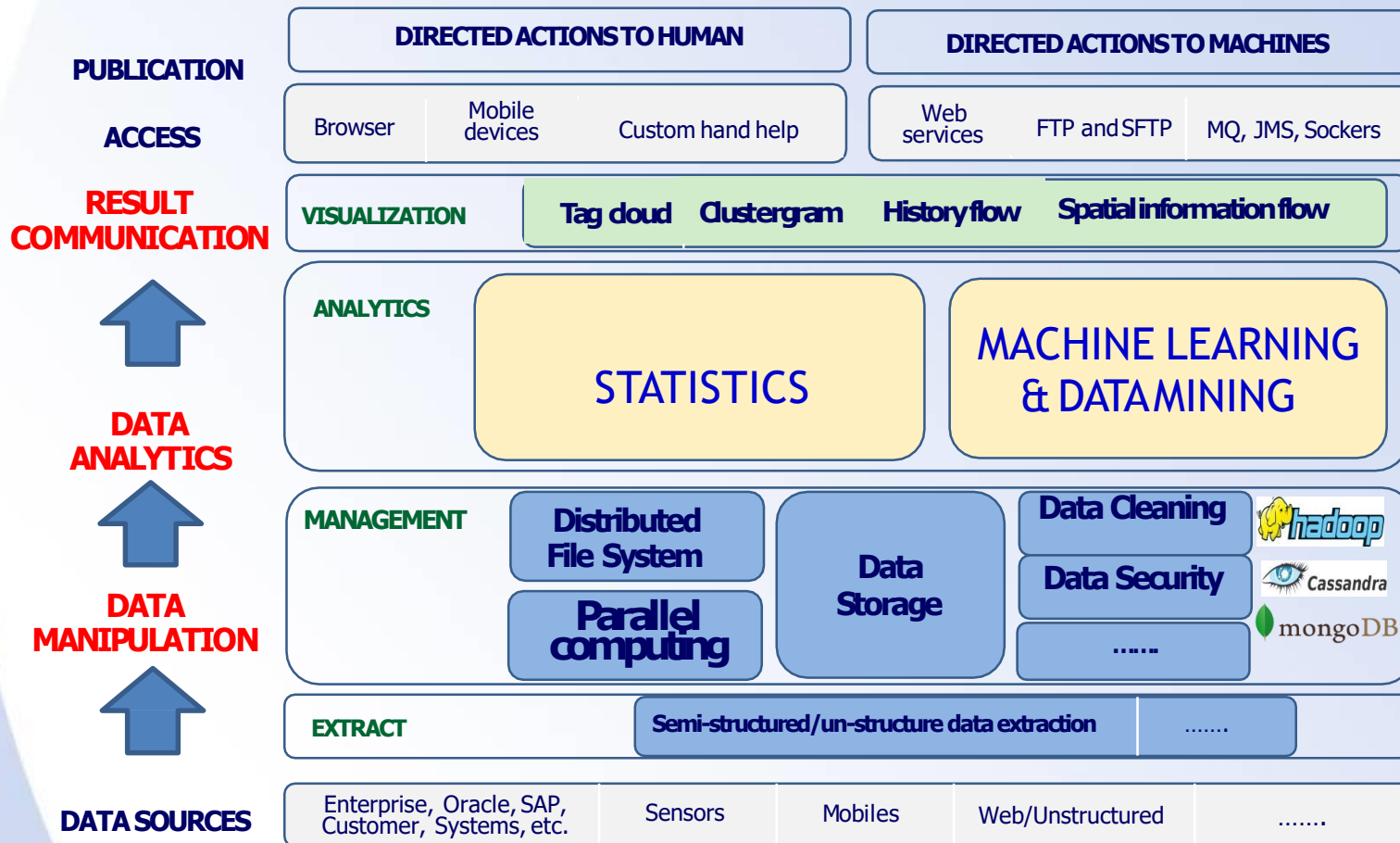


Khoa học dữ liệu là gì



Khoa học dữ liệu là gì

Một lược đồ của khoa học dữ liệu



Source: WAMDM, Web group

Khoa học dữ liệu là gì

Mô hình dữ liệu là gì?

- Ví dụ: bạn muốn dùng một chiếc hộp đen để nhận diện loài vật
 - B1: Bạn phải tìm rất nhiều hình ảnh con chó và con mèo
 - B2: Cho hộp đen đọc những hình ảnh này
 - B3: Dạy cho hộp đen biết đặc điểm nào trên bức hình là của con chó, đặc điểm nào là của con mèo
 - B4: Bạn đưa ra 2 hình ảnh mới, hộp đen sẽ trả lời đâu là hình ảnh con chó, hình ảnh con mèo
- Toàn bộ quá trình này gọi là học máy (machine learning) và cái hộp đen chính là mô hình dữ liệu

Khoa học dữ liệu là gì

- Khái niệm khoa học dữ liệu (KHDL) và quan hệ của KHDL với các đột phá của KH&CN.
- KHDL là nền tảng của công nghệ số, của thời kỳ chuyển đổi số.
- KHDL là một công cụ lao động chính của tương lai. Cần chuẩn bị và trang bị kiến thức.
- Thời chuyển đổi số là cơ hội lớn của phát triển, của con đường mới..

Ứng dụng của khoa học dữ liệu

■ Các bài toán dự báo:

- Dự báo thị trường nhà đất: ngôi nhà ở mảnh đất A liệu có giá bao nhiêu vào năm 2020?
- Dự báo thời tiết: đi nghỉ giỗ tổ và 30/4-1/5 ở Hạ Long có cần mang áo mưa hay không?
- Dự báo hành vi mua hàng: có thích món hàng này hay không? Mức độ thích như thế nào?
- ...

■ Các bài toán ra quyết định:

- Lái xe tự động
- Đặt mua, đặt bán cổ phiếu theo tin tức

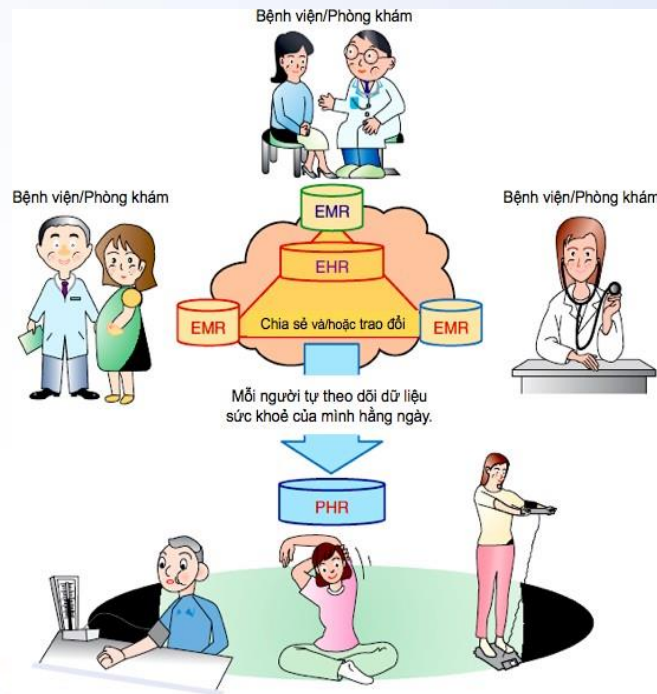
Ứng dụng của khoa học dữ liệu

- Các bài toán ra quyết định:
 - Điều chỉnh nhiệt độ điều hòa tối ưu cho hoạt động của người trong phòng
 - Điều hành xe để đáp ứng nhu cầu của khách gọi taxi
 - ...
- Các hệ thống phân tích thời gian thực:
 - Xu hướng của truyền thông về doanh nghiệp hoặc nhân vật nào đó
 - Cảnh báo cháy qua camera
 - Cảnh báo nguy hiểm với trẻ con, người già
 - ...

Ứng dụng của khoa học dữ liệu

Bệnh án điện tử - nền tảng của e-health

- ❖ **Bệnh án điện tử (BADT, electronic medical records – EMRs) là phiên bản số của bệnh án của mỗi lần nằm viện, tạo và dùng trong từng hệ thống tin bệnh viện (hospital information systems – HIS).**

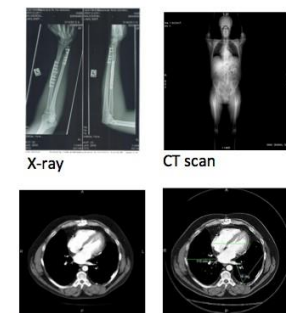


The screenshot shows a medical record system interface. It includes fields for patient information (Name, Age, Gender, Address, Date of Birth, Date of Admission, Date of Discharge, Room, Bed), a list of medical history (Diagnosis, Date, Time, Doctor, Location), and a list of medications (Medication, Dosage, Frequency, Route, Start Date, End Date, Status). The interface is in Vietnamese and includes a search bar and a list of filters.

Nhận xét và ý lệnh hàng ngày của bác sĩ và ghi chép thường xuyên của y tá về tình trạng của người bệnh.

DỮ LIỆU LÂM SÀNG (VĂN BẢN)

DỮ LIỆU CẬN LÂM SÀNG (SỐ & ẢNH)



PACS (Picture Archiving and Communication System)

Dữ liệu xét nghiệm (máu, nước tiểu, điện tâm đồ...)

Yasuo Ishigure, Trends, Standardization, and Interoperability of Healthcare Information, NTT Technical Review 2017

Ứng dụng của khoa học dữ liệu

Heterogeneous and longitudinal data

Electronic medical records (EMR)

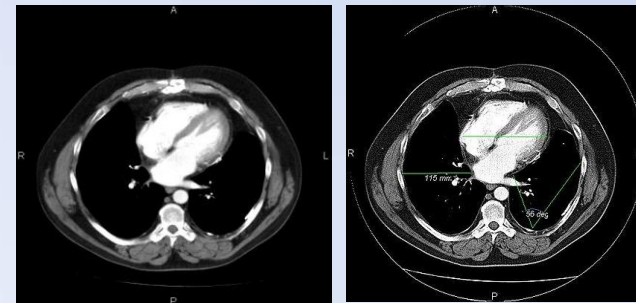
984,20123,1216,0,3354-02-05 05:40:00 EST,3354-02-05 06:01:00
EST,4270,"N",54,"Nursing/Other", "MICU nursing admission note 7AM",
MICU nursing admission note 7AM

Pt is 68 yo male adm [**Hospital1 2**] EW [**2-4**] s/p fall 2 weeks ago while in [**State 552**] where he landed on his left side, having left sided pain. CXR x 2 at hospital, no fx, pt sent home. Took motrin for pain steadily last 2 weeks. ^ SOB, anorexia last 2-3 days. Flew to [**Location (un) 175**] for medical care. In EW, + EKG changes, + troponin/MB. ARF, cr 3.4, K 5.4. Given Kayexelate, D50, IV insulin, CaGluc. Heparin gtt started for EKG changes, ?PE. No CT d/t ARF. VQ scan showed low prob PE. Also FSBS 300s, covered by SQ insulin. Vanco/levoquin for ? UTI. Desattened on RA, 100% NRB with SATs 100%. CXR no rib Fx. Bicarb gtt for acidosis, gap 26. Hemodynamically stable, BP decreased 80s x 2 while sleeping, increased when awake. A&O x 3. Tx MICU for further management. ARF probably d/t motrin use.

Neuro - A&O x 3. C/O left sided pain when turning, otherwise comfortable. MAE.

Resp - Weaned O2 NC 6L, SATs 94%. Lungs clear, diminished at bases. No SOB.

CV - BP 103-118/54-59. NSR 70s-80s, no ectopy. Heparin gtt 1450U/hr. PTT >150, shut off @ 4:30. Restarted 6:00 @ 1200U/hr. K 5.7->6.5. EKG unchanged. 2amps CaGluc, insulin 10U IV, 30gm Kayexelate given. Pt has had no stool from any kayexelate given. 4:30 lytes will not reflect



X-ray, CT scan, MRI

MCHC	327.0	g/L	280 - 360	280 - 360	06/10/2016 14:5
MCV	81.2	fL	83.0 - 98.0	83.0 - 98.0	06/10/2016 14:5
MPV	9.6	fL	6.0 - 13.0	6.0 - 13.0	06/10/2016 14:5
Mid#	1.5	GPL	0.2-0.8	0.2-0.8	06/10/2016 14:5
Mid%	21.9	%	5 - 8	5 - 8	06/10/2016 14:5
P-LCR	22.2	%			06/10/2016 14:5
PDW	11.3	fL	6.0 - 10.0	6.0 - 10.0	06/10/2016 14:5
RBC(Hồng cầu)	4.67	mm ³	4.0 - 5.9	4.0 - 5.9	06/10/2016 14:5
RDW	40.1	%	8.0 - 12.0	8.0 - 12.0	06/10/2016 14:5
THR(Tiểu cầu)	238	mm ³	150 - 450	150 - 450	06/10/2016 14:5
WBC(Bạch cầu)	6.9	mm ³	4.0 - 10.0	4.0 - 10.0	06/10/2016 14:5
Tổng protein nước tiểu (B)	7.0		4.8-7.4	4.8-7.4	06/10/2016 14:5
pH					06/10/2016 14:5
BIL (Bilirubin)	Ẩm tính	umol/L	<3.4	<3.4	06/10/2016 14:5
BLO (Hồng cầu)	VẾT	/μ	<5	<5	06/10/2016 14:5
GLU (Glucose nước tiểu)	Ẩm tính	mmol/L	3.7 - 6.2	3.7 - 6.2	06/10/2016 14:5
KET (Ketone)	Ẩm tính	mmol/L	<5	<5	06/10/2016 14:5
LEU (Bạch cầu)	-	/μ	<10	<10	06/10/2016 14:5

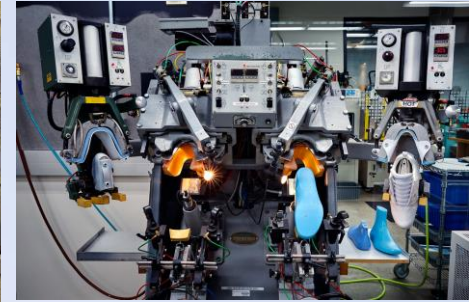
Lab examination (blood, cardiogram...)

CLINICAL DATA (clinicaltext)

PARA-CLINICAL DATA (numbers)

Ứng dụng của khoa học dữ liệu

May mặc, đóng giày sẽ tiến đến đặc chế



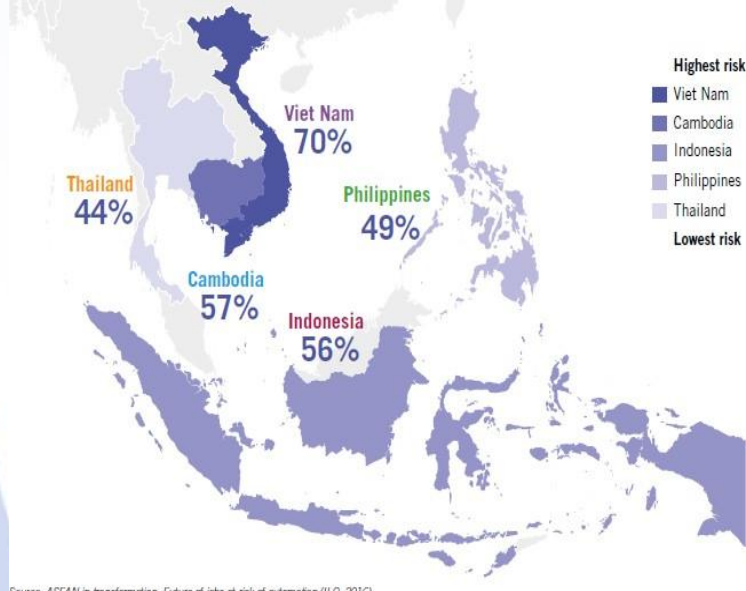
Điều gì đang xảy ra? Sao đặc chế tự động được?

Ứng dụng của khoa học dữ liệu

Việc dễ tổn thương trong chuyển đổi số

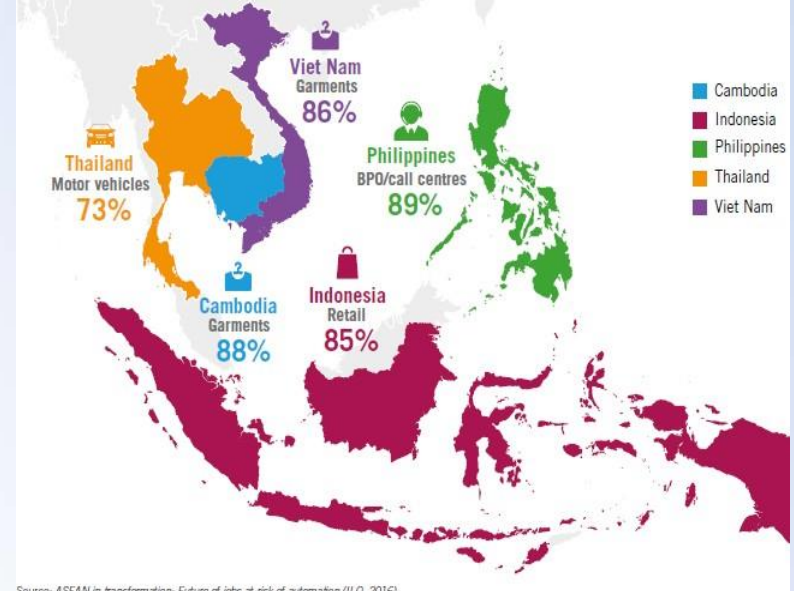
WORKERS AT HIGH RISK OF AUTOMATION IN ASEAN-5

Simulations on jobs susceptible to digitalization



PERCENTAGE OF WAGE WORKERS AT HIGH RISK OF AUTOMATION IN KEY SECTORS IN ASEAN-5

Simulations on jobs susceptible to digitalization



ILO: ASEAN in transformation: the future of jobs at risk of automation (2016)

Ứng dụng của khoa học dữ liệu

1. Hãy nêu một vài vấn đề liên quan đến địa phương (quê) của bạn, mà bạn cho rằng có thể giải quyết bằng khoa học dữ liệu.
2. Theo bạn có những vấn đề nào của trường ta có thể là đối tượng nghiên cứu của khoa học dữ liệu?
3. Gần đây Facebook có vụ bê bối vì lộ thông tin cá nhân của khách hàng, bạn có cho rằng các thông tin mà bạn đưa lên facebook là quan trọng?
4. (vui) Đánh số đề có phải là bài toán của ngành khoa học dữ liệu?

Nhà khoa học dữ liệu (Data scientist)

- Với skillset chuyên sâu và trải dài trên nhiều lĩnh vực
 - Math and Statistics
 - Programming and Database
 - Communication and Visualization
 - Domain Knowledge and Soft Skills



Nhà khoa học dữ liệu làm gì?

- Thu thập và xử lý dữ liệu để tìm ra những “insight” (giá trị bên trong)
 - Ví dụ: dựa trên các thông tin thu thập được từ các post/comment/status trên mạng xã hội, Data Scientist có thể tìm ra được: cứ gần đến ngày valentine thì tần suất xuất hiện các thương hiệu ABC cao hơn hẳn
- Giải thích, trình bày những insight đó cho các bên liên quan, để chuyển hóa insight thành hành động
 - Ví dụ: khi tìm ra được insight giá trị từ data, bạn cần làm report/presentation hay visualization để biểu diễn, giải thích cho các bên liên quan hiểu được

Việc mới và kỹ năng mới?

- Khi sản xuất thông minh phát triển, nhiều loại lao động tăng lên (liên quan phân tích dữ liệu), nhiều loại giảm đi, nhiều loại lao động mới xuất hiện, dù chưa biết.
- Văn hoá nghề nghiệp thay đổi, chuẩn mực và đạo đức nghề nghiệp mới
- Người lao động cần nhiều kỹ năng mới: kỹ năng quản lý, kỹ năng kỹ thuật (công nghệ số), kỹ năng mềm.
- Để đào tạo nguồn nhân lực số, cần thay đổi đào tạo: mục tiêu, nội dung, cách thức.

Các yếu tố phát triển khoa học dữ liệu

Sản phẩm data là gì?

- Sản phẩm data được xây dựng dựa trên dữ liệu
 - Tính năng recommendation của Amazon được xây dựng dựa trên dữ liệu của nó: người dùng muốn mua món đồ gì? Những món đồ nào nên mua kèm?
- Sản phẩm data có thể là một sản phẩm riêng biệt hoặc một phần trong sản phẩm lớn
 - Facebook có thể tự tag ảnh bạn bè của bạn
- Sản phẩm data bao gồm nhiều thành phần nhưng mô hình dữ liệu là cốt lõi của nó và được xây dựng bằng các thuật toán học máy

Các yếu tố phát triển khoa học dữ liệu

Phân tích S W O T về khoa học dữ liệu của ta

ĐIỂM MẠNH (STRENGTH)

- Có kinh nghiệm đào tạo toán
- Đông đảo người làm CNTT
- Tuổi trẻ và khát vọng
- Nhiều truyền thống dân tộc

ĐIỂM YẾU (WEAKNESS)

- Ít dữ liệu cần cho các ngành nghề
- Thống kê toán và ứng dụng toán
- Đặt bài toán có ý nghĩa
- Tính thích nghi chưa cao
- Tiếng Anh còn hạn chế

- Phát triển nhiều ngành nghề với tiến bộ của công nghệ số
- Kết hợp nhà nước-doanh nghiệp-trường viện
- Kết hợp trong và ngoài nước
- Nhiều loại công việc mới

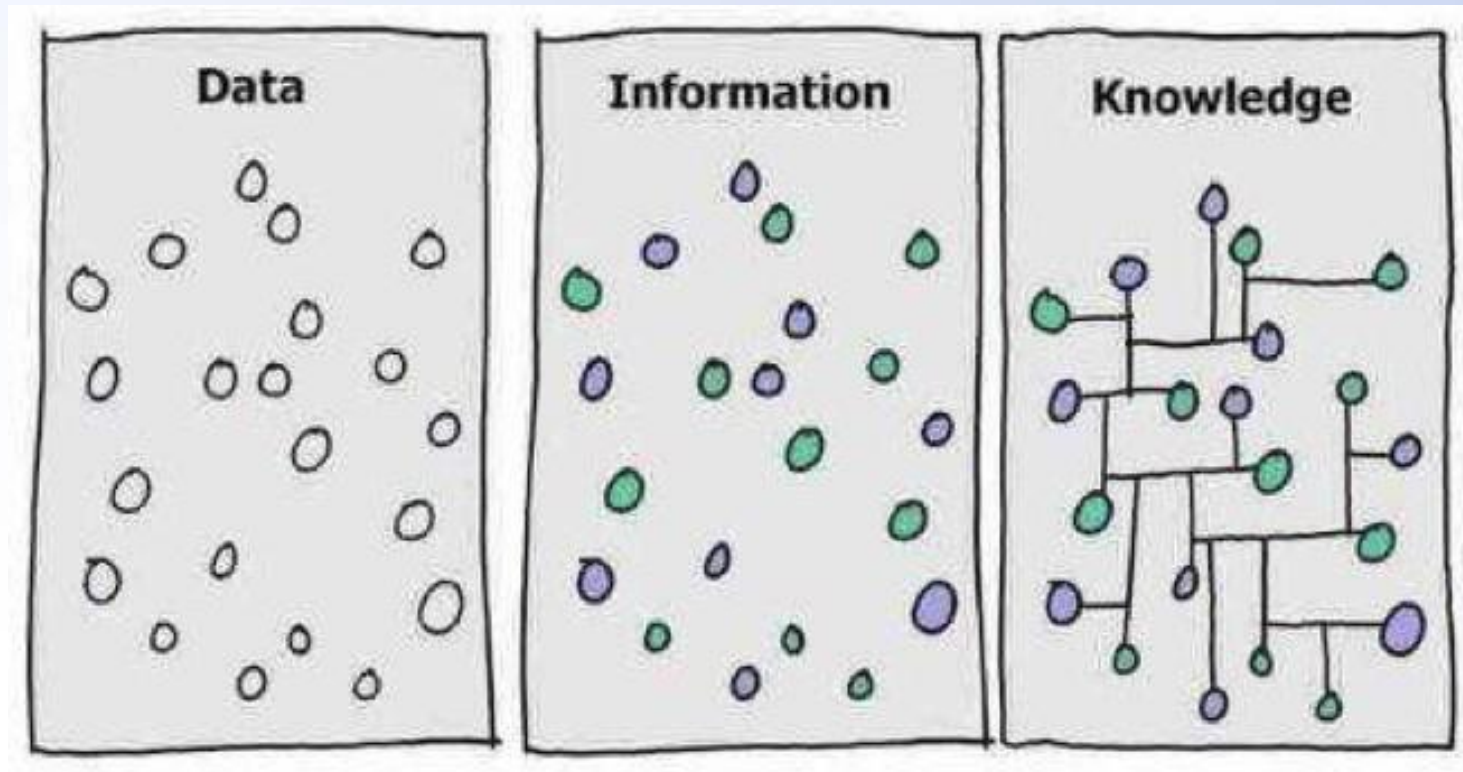
- Tư duy mới và sáng tạo
- Khơi dậy được truyền thống và sức mạnh dân tộc, đặc biệt với tuổi trẻ
- Chiến lược, chính sách đúng đắn
- Đào tạo nhân lực công nghệ số
- Tạo ra nguồn dữ liệu
- Đưa KHDL vào mọi ngành nghề

CƠ HỘI (OPPORTUNITY)

THÁCH THỨC (THREAT)

Các yếu tố phát triển khoa học dữ liệu

Data, Information, Knowledge



Data scientist workflow



Data scientist workflow – Bước 1

■ Input

- Workflow bắt đầu từ một yêu cầu hoặc nhiệm vụ: “Nhu cầu tìm kiếm hình ảnh của Google: đưa cho máy 1 bức ảnh, trả về những bức ảnh tương tự”
- Nhu cầu này có thể bắt nguồn từ:
 - Do bộ phận business thu thập phản hồi từ người dùng và đề nghị có thêm tính năng ABC
 - Hoặc, do chính Data Scientist khi làm việc với dữ liệu, nghiên cứu đặc tính của sản phẩm/ công ty cũng như kiểu/ lượng data hiện có...thì nảy sinh thêm sáng kiến phát minh tính năng XYZ

Các yếu tố phát triển khoa học dữ liệu

Data scientist workflow – Bước 2

■ Lên kế hoạch

- Làm tính năng này có khả thi hay ko?
- Sẽ cần loại dữ liệu gì? Ở đâu? Bao nhiêu là đủ? Lấy dữ liệu như thế nào?
- Cần bao nhiêu resource (nhân lực, thời gian)
- Tính năng này sẽ được gắn vào đâu trong sản phẩm cuối cùng và sẽ giúp ích được gì cho người dùng

Data scientist workflow – Bước 3

■ Thu thập và làm sạch dữ liệu

- Để dạy cho máy biết phân biệt chó/mèo, nó càng phải học nhiều hình ảnh càng tốt. Nên phải đi “gom dữ liệu”
- Dữ liệu gom xong sẽ còn lộn xộn và nhiều rác thì phải “làm sạch dữ liệu”.
 - Hình ảnh ko cần thì loại bỏ; Hình mờ thì làm cho rõ ...
 - Đồng bộ hóa dữ liệu
 - Hình ảnh mang về có kích thước khác nhau, phải đưa hết về cùng kích thước, định dạng theo mô hình dữ liệu đã chọn
- Nếu dữ liệu chưa đủ phải thu thập thêm

Data scientist workflow – Bước 4

■ Chọn giải pháp

- Nếu vấn đề đã có sẵn giải pháp
 - Lựa chọn / kết hợp các giải pháp lại, chạy thử nghiệm, kiểm tra thử nghiệm nào tốt nhất và vì sao, chọn giải pháp để phát triển thêm
- Nếu vấn đề chưa có sẵn giải pháp
 - Cần làm nghiên cứu: tìm hiểu xem trước mình đã có ai từng làm về vấn đề này hay chưa
 - Sau đó, chọn ra một hoặc một loạt các phương pháp để thử nghiệm

Data scientist workflow – Bước 5

■ Máy học

- Chạy thử mô hình và đánh giá hiệu năng
 - Tưởng tượng bạn điều khiển bằng điều khiển với nhiều nút. Bạn thử chỉnh nút này 1 chút, thấy kết quả ra tốt hơn chút xíu thì giữ lại và chỉnh thử nút khác
- Nhận diện các yếu tố ảnh hưởng đến kết quả. Điều chỉnh dấu hiệu ưu tiên để ra được kết quả tốt nhất.

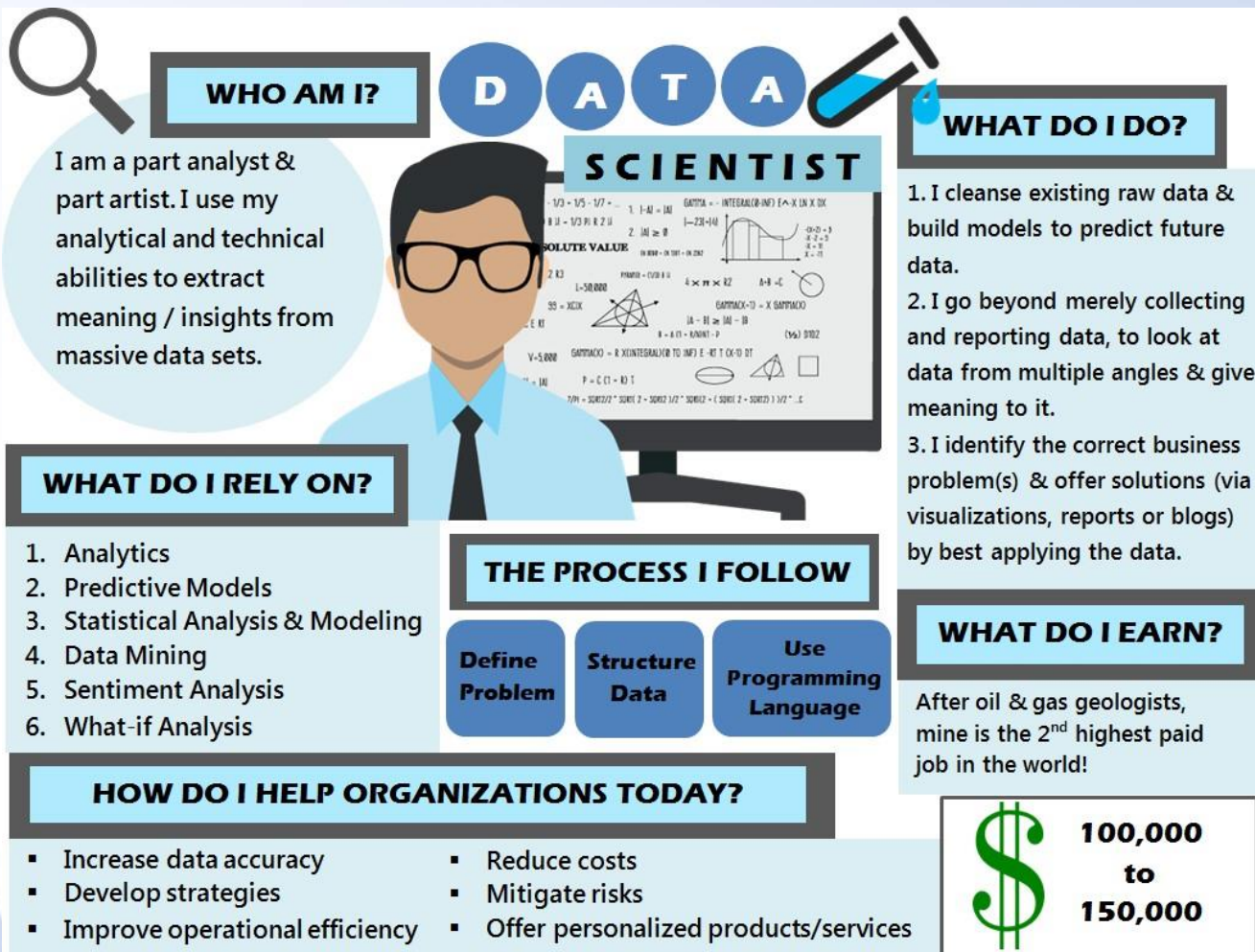
Data scientist workflow – Bước 6

■ Output

- Kết quả gắn vào một sản phẩm lớn có tính ứng dụng
- Viết bài báo
- Tổ chức hội thảo

Các yếu tố phát triển khoa học dữ liệu

Data scientist: tổ chất cần có?



Data scientist: tố chất cần có?

■ Kiên nhẫn

- Tố chất này cực kì quan trọng vì DS phải dành phần lớn thời gian để thu thập và làm sạch dữ liệu

Ví dụ, bạn muốn làm một model dự đoán giá nhà.

Bạn sẽ phải thu thập dữ liệu về nhà từ nhiều nguồn khác nhau.

Mỗi nguồn này lại lưu dữ liệu theo một cấu trúc riêng. Vậy thì bạn phải quy chúng về một cấu trúc chung.

Sau đó, bạn làm sạch bằng cách loại bỏ các dữ liệu không phù hợp, như:

- *Dữ liệu thiếu: có số lượng phòng mà không có diện tích.*
- *Dữ liệu rác: diện tích 10m2 mà giá 200 tỷ.*

Data scientist: tố chất cần có?

■ Giao tiếp tốt

- Với Team Business: để hiểu rõ hơn về sản phẩm cũng như requirements, từ đó tìm ra các insights có giá trị
- Với Team Engineer: để áp dụng mô hình của mình vào hệ thống hoặc đề nghị họ tổ chức/hệ thống data cho mình sử dụng
- Trình bày, giải thích insights cho các bên liên quan hiểu

Các yếu tố phát triển khoa học dữ liệu

Data scientist cần kỹ năng gì?

■ Kiến thức toán học: yếu tố quan trọng số 1

- Nghề data science sử dụng nhiều kiến thức liên ngành.
 - Machine learning là sự kết hợp của các mô hình toán học chạy bên dưới
 - Khi xử lý / làm việc với dữ liệu, bạn sẽ cần sử dụng rất nhiều kiến thức về toán, xác suất thống kê,...
 - Tư duy toán học sẽ giúp bạn dễ tiếp thu và học các kỹ năng khác nhau

Ví dụ, khi cho máy học một bức ảnh để phân biệt con chó với con mèo. Thì bức ảnh đó sẽ được chia làm nhiều vùng tương ứng với 100 ô vuông chẳng hạn.

Rồi bạn dạy cho cái máy rằng, trong bức ảnh, ô ở cụm phía bên góc trái có nhiều màu đen, kết hợp với ô ở cụm phía bên góc phải có nhiều màu trắng, thì đó là đặc điểm nhận biết con chó.

Data scientist cần kỹ năng gì?

- **Khả năng Lập trình phần mềm**
 - Công việc của Data Scientist rất gần với Software Engineer. Vì vậy, code cũng là một yêu cầu quan trọng
- **Sự nhạy bén**
 - Khi nhìn vào dữ liệu, bạn cần đủ nhạy để suy đoán: đối với loại dữ liệu này thì nên làm gì với nó, nên estimate như thế nào?
 - Sự nhạy bén là tố chất song cũng tích lũy dần theo kinh nghiệm và thời gian

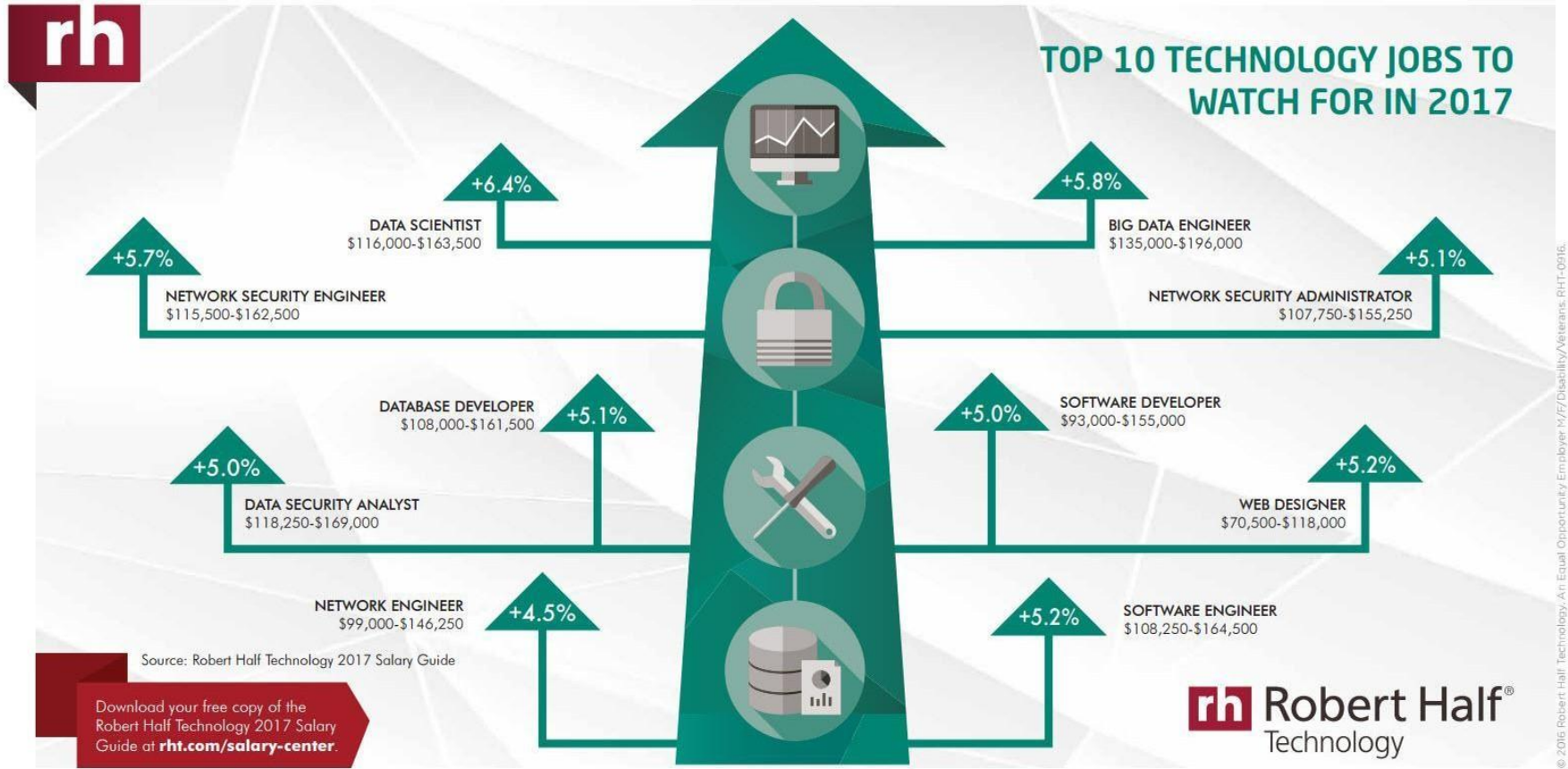
Nhà khoa học dữ liệu làm gì?

Data scientist cần kỹ năng gì?

- Nghề Data Scientist đòi hỏi khá nhiều kiến thức và kỹ năng tổng hợp
 - **Machine Learning**: để học từ dữ liệu, từ đó tạo ra các mô hình dự đoán
 - **Database**: giúp lưu trữ, truy xuất dữ liệu cũng như thực hiện tính toán
 - **Programming language**: viết code để áp dụng các mô hình đã học được nói trên vào sản phẩm cụ thể hoặc để thao tác với database
 - **Visualization**: giúp hiểu hơn về dữ liệu hoặc trình bày kết quả phân tích

Các yếu tố phát triển khoa học dữ liệu

Nhu cầu



Các yếu tố phát triển khoa học dữ liệu

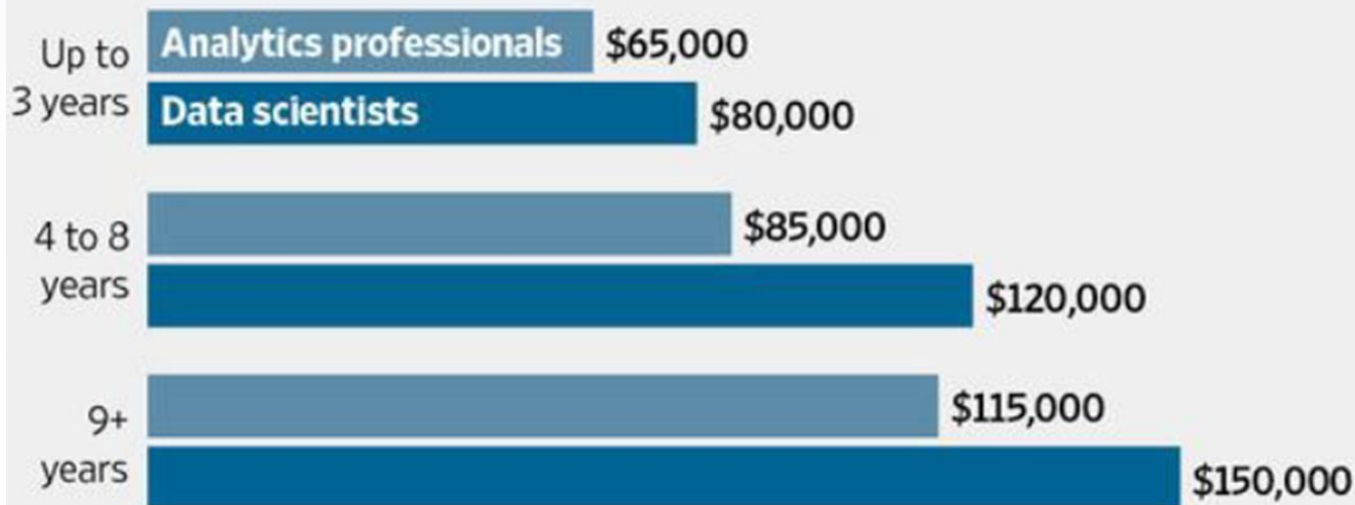
Nhu cầu



Thu nhập

Big Data, Big Paycheck

Median salary for analytics professionals and those specifically within data science, by level of experience.



Note: Data do not include managers Source: Burtch Works

The Wall Street Journal

LOGO

CẢM ƠN!