



# Fostering Big Data Applications in HK enterprises and Strengthening Collaboration of Digital Economy in the Greater Bay Area

- 公民數據科學家 Citizen Data Scientist
- 大數據潛力與數據經濟 Big data potentials and data economy
- 大數據治理概述 Overview of big data governance
- 數據治理原則 Principal of data governance
- 人工智能 Artificial Intelligence
- 機器學習 Machine Learning
- 深度學習 Deep Learning
- 商業建模 --- Business Modelling & Cross Border Data
- 數據應用 -- Big Data Application Health & others
- The future of ecosystem of data economy -- Panel

### 1<sup>st</sup> Session -- Oct 8, 2021



#### **Data Governance for the Digital Economy**

#### Ir. Allen Yeung, JP Founding Chairman, Institute of Big Data Governance

Oct 08, 2021

### 2<sup>nd</sup> Session -- Nov 12, 2021



#### Agenda

- Part I : About iBDG
- Part II : Data Governance Why Do We Care? Regulations Around the World
- Part III A : Data Governance Best Practice and Independent Verifications Who Needs to Do What?
- ▶ Part III B : iBDG Big Data Governance Best Practice and Certification Programme
- ▶ Part IV : How iBDG Big Data Governance Certification Aligns with Regulations
- Part V : Questions & Answers Session
- ▶ Appendix : iBDG Data Governance Assessment Criteria A Closer Look

Page 3



### 3<sup>rd</sup> Session -- Dec 17, 2021

- Use of Big Data
- Cloud Computing
- Big Data Recent Trends
- Artificial Intelligence
- Example of CBA -- Tiktok

Charleston Sin Executive Director, MIT HK Innovation Node Secretaire General, IBDG

### Nowadays ... If I sell shoes ...





### If I sell shoes...





Take them off

need it for work

More comfortable pls

# If I sell shoes...



- CEO/CXO -- strategize data
  - IT Engineers -- design the system, & develop app
  - Data scientists -- analyze data
  - Designers -- understand and use data
  - Marketing -- understand data and algorithms and use data
- Shop assistants -- use the system

### **NETFLIX: DATA-INFORMED TV STREAMING**

Location: Los Gatos, Calif.

**How it's using big data**: The premise of Netflix's first original TV show – the David Fincher-directed political thriller *House of Cards* – had its roots in big data. Netflix invested \$100 million in the first two seasons of the show, which premiered in 2013, because consumers who watched the British series *House of Cards* also watched movies directed by David Fincher and starring Kevin Spacey. Executives correctly predicted that a series combining all three would be a hit.

Now, eight years later, big data impacts not only which series Netflix invests in, but how those series are presented to subscribers. Viewing histories, including the points at which users hit pause in any given show, reportedly influence everything from the thumbnails that appear on their homepages to the contents of the "Popular on Netflix" section.









t=1

t=2

t=now

MEYER

# **Disruptive Digital Technologies**



# Demand Side SME, Enterprise, you & me



# STARTING ECOMMERCE BUSINESS IN HONG KONG

# Supply Side Cloud Service Provider

- pull in computing resources to serve multi-user simultaneous
- backup with redundancy
- resource is shared





# 规模的力量 ... 互联网+海缆+全球云数据中心

Cloud



是高性能計算,例如藥物發現、基因組等…降低初創企業的准入門檻



大數據

# Big Data describes Tools, Technologies and Practices to process and analyse massive dataset.



- Velocity
- Varity
- Veracity (quality)
- Value





Data Mining

# 数据分析的趋势





- 代理数据的创造性使用 (Creative use of proxy data)
- 数据可以交换价值 (Data can be exchange for value)
- 汇总数据--来自不太可能来源
- 数据时代与无现金时代
- 跨境数据



# #1 -- Creative Use of Data











大數據

"Octopus data is one of the important data sources and could help HKUMed identify the contact patterns and density of city residents in different areas and could play a significant role in combating the pandemic," said Professor Gabriel Leung.

Mobile Data => Social Economic Class

http://senseable.mit.edu/singapore-calling/

Octopus Data => COVID-19 diffusion before announcement

# #2 -- The Value of Data 數據資產化 ... Data Exchange Economy



大數據

# 2020 李小加倡提速培育數據要素市場

身兼全國政協委員的港交所(00388 行政總裁李小加【圖】在兩會提案,建議 由中央深改委牽頭,由國家發改委組織 實施,可從醫療數據着手,在國家衛健 委的組織下,與業界共同制定相應監管 規定,以具有前瞻性和包容性的視角,來 加速數據要素市場培育的進程。

李小加指出,在此過程中,應盡早聚焦幾 個核心問題,一是數據的確權問題;二是數據的安全 和隱私保護問題;三是數據的分享問題;四是數據 市場的監管問題;五是數據的交易問題。 他認為,由於資料不同於其他生產要 素,需要釐清的問題無窮無盡,建議在初 始階段聚焦醫療健康領域並以試點實 行,因為醫療大健康基於以下的特性最 有可能實現突破。

#### 醫療健康領域先設試點

首先是價值最大,醫療數據與人類健康密 切相關,可視為重要的戰略資源:其次是最迫切,生 物科技的發展急需醫療健康數據的支援,新冠肺炎 疫情的防控工作使得犧牲部分隱私和訊息保護權利 來維護公共衛生安全成為可能:三是最容易落地,因 為醫療訊息的歸屬相對清晰,只要做好核心數據的 保護工作和解決使用權計價問題,就能馬上開展應 用:四是最能產生經濟效益,在隱私加密保護的前提 下,也許政府可以收取部分醫療數據市場化帶來的 增值税,用以補貼醫保支出。

李小加表示,在新時代,借力新要素,中國可以 憑藉強大的數據積累和技術軟實力,為下一輪世界 經濟發展輸出「數據能源」。同時,在全球聯手抗疫 的今天,加速醫療健康領域大數據的共用和互助、共 同維護公共衛生安全,更具有緊迫性與現實意義。

第三是數據分享問題:政府及企業應考慮是否 可以或如何分享數據。李小加舉例,政府通過公共 服務產生的數據,在不涉及安全及隱私的情況下是 否應允許社會共享?政府在甚麼情況下可用其控制 的數據參與交易,並獲取收入?如何正視互聯網、 金融與電訊服務等平台產生的海量數據已開始大量 商用這一現實?

第四是數據市場的監管問題:監管應訂明全國 統一的「數據禁區」,列明不能共享或交易數據, 同時設置數據安全和私隱保護方面的標準和要求。

HEYT	
資本巨頭扔 <b>垂涎背後</b>	
新式茶飲店在全國運地開花, 競爭異常激烈,但不少資本巨頭仍 然類意測入這條擠逼的「聲道」。有 內地輿論認為,這些資本巨頭看好 的不確僅是眼前的茶飲品牌,更多 的是品牌背後的消費大數據,以及 數據背後的年輕族群。	地F幹財經」引述中國食品產業分析 師朱丹蓬觀。 走進一家喜茶的門店,幾乎看 不到人排隊,如果你走到收銀台前 準備下單,店員會推薦你用小程序 錢上點單,並說明你大概需要等待 多長時間,前方還有多少杯正在排
消費數據有助推出新品 有數據顯示,今年上半年內地 茶飲品牌的融資項目有十五宗,披 當金額超過五十億元人民幣,披露 金額已高於去年全年。其中著名品 牌寫茶上周剛完成了五億美元(近	隊製作中。這樣的場景在許多新式 茶飲門店都普遍存在。 分析認為,數碼化已經是新式 茶飲的標準配置,資本巨頭看中 的就是新式茶飲的數碼化能力,而 非單純判斷[現在職不賺錢]。相關 產品銷量和地區分布等消費數據報
附著来上,所附无成, 1 1.8 天/1 1.0 四十億港元)的融資, 投資方有紅 杉資本,高瓴資本與聽訊資本等眾 多知名機構。 「新式茶飲最大的價值是背後 的流量、大數據與用戶畫像。」內	律前加重和這個方面等的有效還得報 告,將幫助品牌了解用戶裏好與需 求,更精準、快速地推出新品。這 些數據無論是對做新式茶飲,或者 拓展其他餐飲甚至生活服務領域, 都具有長期價值。



HONG KONG MONETARY AUTHORITY 香港金融管理局

# Fintech 2025



# #3 – The Value of Aggregated Data



Commercial Data Interexchange (CDI)

**BIG DATA** 







規收集及使用個人訊息·遭中國監管機構勒 令把 App 下架。

李小加認為,數據安全涉及私權及公權 兩方面·內地政府除擔心掌握大量個人數據 的大型企業會出現濫用之外,在地緣政治緊 張下,更視數據涉及國家安全,「誰知道將 來兩國有敵對關係時·這些數據有沒有可能 成為新的情報來源。|

地

數

據

監

管

涉

或

安

中國無現金社會可領導全球 中美關係緊張·李小加指出·中國於過 去40年來一直在政經系統中靠攏西方,但 不想再互相依靠。他又提到·目前 界經濟體系中仍處一統天下的角

色·若中國整個經濟能率先進入無現金社會 的數碼大潮·有機會改變現行的規則·帶領 全球。 李小加補充,香港應把握

的浪潮,助中國的中小微企引入國際資金。 攝不應做「攔路虎」,應允許試錯 及創新、市場自會找到出路。



李小加:無現金時代將臨

DIDI) 继赴美下市:

· 將會是一個很大的挑戰

#### 未來投資重點 將放在小微企

李小加在出席一個論壇時表示·未來數

排時代中·最先完成改革的國家。€

# #4 – Data & Cashless Era



- 未來金融發展的新浪潮就是無現金時代的到來,香港亦需要抓住這一 浪潮,才可以帶領世界金融市場的發展。
- 但如果中國能夠率先進入無現金的互聯網數字化時代,未來在金融領 域中,新的規則可能不再由西方統領。
- 他又指,由於中國過往的傳統服務業做得太差,長期未能滿足人民生 活需求,因此數字化發展過程中沒有既得利益者的阻撓,因此其數字 化發展更加快,更有機會成為數據時代中,最先完成改革的國家。

https://www.ourhkfoundation.org.hk/en/event/90/our-hong-kong-foundation/hongkong%E2%80%99s-role-and-opportunities-sustainable-financial

Sources: OHKF, HKET, HKEJ -- July 6 & 7 2021



# **#5 -- Cross Border Privacy Data** 特事特辦,非常時期用非常手段



145 'भर' 	HA Go	ан э <b>—</b> 92	7:18	at 🗢 🗖
Alfraic	( 「お有HA746		<b>く</b> 我照顧的人	照顧我的人
			我照顧的人 🕀	1位成員 ~
		roor Gebridderi	×xx	
********	) #fate	-#12	最多可邀請十人,並定	明檢視關係
Radu	(+)		我的子女 🕣	0位成員 >
		$\odot$		





### 大數據

Source: HKET HKEJ

1.6

5 9

2.5

-----

# iBDG



# #5 – Green Channel & Data Capsule

# 倡建港深數據專區



作為最先為之京、 海敷 該書面提供建築及社動者之 一,筆者更得有必要在這個時 使提出建立海家戰隊專單的必 要性和機會,從而加速大灣區

早至2015年8月,國家就創布了《伊維大數 據發展行動詞要》,写在率先推動數字產業化的建 設。並於同年9月在貴州戲頭。全國首個大數觸 如合就驗盘的建設1件。至2016年10月,第二 批獲比據設備家級大數漆約合試驗量的省份名單 山合,其中包括兩個綺密域顯綜合試驗品:《淨 第一版正三角別。

這些試驗區的建立(一個共通性,主要是打 加區內的被字查案以加進新經濟稅底,從成效火 看,「試驗值 幫助了部分的科技創新,但真正的 可找備數字时代達未來證,我們需要有更多網驗的 數據安全保護措施,個人數據私隱保護方式,以及 企業之間在數據共早上的規管。

2019年2月千餐的《粤港澳大迎區發展氣 劃調要》,提出了大灣區在藝處產如一的損居方 向,逆見是世一年儿的難識,因為莎及法律、文 化,治理規管方面的產用,這成了開或內的數據資 源積度分化;而是,商業活動上的需求又告訴我

化:大鸡直的評論協同有解於三流(人流、物流、资 計添)互聯互通,近期深渗合作已經成為大導面的 核心方向,而两市之間的數據打通尤其重要,也可 望為區內其他,9+2城市|起到示的作用。

有谋於比,至者認為有必要深入研究違立港 深設處本島的可行性,宜先從實業為行的地方若 手,抒機制與系統建立起來,起然各地的數據等區 有若不同的挑戰,当者勿望在這裏狀碍引下,分子 北京在被據書區方面的部分經驗,具中有《開於通 這公共數據開放征證人,1智能產業發示的,1件方 策》的穴個見路。

 1)宣施分級分類管理,保障公共数據開放有 庁宮施:

2)深入推進一般公共數據開放,為人工智能 蓋美發展得供普惠歐要供給; 3)建設公共數據開放創新基地,頒過与定方

式向人工智能企業有條件對放數據; 4) 落實在城市管理領域開展人工智能應用;

5)支持引導公共服務領域問題人工治能應用: 6)持建人工智能生態增紧, 打造人工智能大 致機健康發度倚誤境。

#### 車品費 紅杉資本中國專家合夥人, 阿里巴巴商學院特聘教授暨學術委員會委員



4 前工值投情,收到涂 圳市收款元的泡渍作為「改樣 權」研討會写家之一,曾百己 完民大數據 太實義。最近《深圳經濟行區 改讓幾續例》《、相《條例》》獲市一屆人大常委會第二 二次會議通紙,2022年1月1日開始實施《條例》 的內容沒為了個人數錄,公共數項,歐張生產要素 市場,或處安全等方面,是國內或求領域首部基礎 性,統合件立人法案。

其中數據確碰問題一百幅受評議,正方觀點 認為該據是在用戶使月過程中產生的,該據的三人 理應是川戶率人,而不是企業;反方覺得該該並非 用戶單獨的行為所筆生,而是因為取務商提供服務 而引致或據的產生。其於1公領公有理,要試婆有 理,在所有數據權的問題上,理應由企業和用戶 案商辦決。

在2018年,杭州互聯鎖法院的數據產品不正 常競學結紛奏的首宗觀於茲處確權判例,原告淘賞 公司在其半台所收集的用戶瀏覽,搜案,收藏、交 易等行為數據的基礎上,而為約實所開發的线計運 營分所紧続。

#### 检告B公司以提供還程登錄的方式,幫助他人 覆取涉案或據產品中的数據內容,並從中後取利 益,到间中运官高周涉索数據產品是消費公司付出 的人力,物力,吃力,是經過長期經營情累而成,確 諭了 企業對具投入大量智力勞動成果形成的數據 產品和服務具有影產性增益」,因此渴實公司對其 數據享有經營及財產權益。 由此可見,景定數據確確的目的不但是確保

五式可元 57% 就發電價(10)(57)。但足幅(5 数據的來原有俗源,這可以對茲據描述出去的價值 進行確認,也可以使數據分發,獲取和使用的邊养 更清晰,更有利於數據資源在企業中進行流嗎。

《條例》確立自然人對個人滅滅依法享有人容 構益,包括知情同意,補充更正,制除,查問,複製 等權益:同時自然人,法人和非法人組織對合法處 評數擬形成的數據產品和服務享有法律,行政法規 及本條例規定的財產權益,可以依法自主使用。 筆者認為僅有在數處權益清晰驚定的情況。

至47%為僅有什麼像產品內喻量定的有近 下,數字經濟才可以得到體康發展。香港作為深圳 的鄰居,應該总索如何配合得宜,也可能會成為一 個新的機道。 車品覺

紅杉資本中國專家合夥人 阿里巴巴商學院特聘教授暨學術委員會委員

### 數據流通機制成契機

中國正式通過《數據安全 8 法》為維護畝被安全領域的基 碰性法律及保障数法安全。数 安法明確要求建立數據分級分 全民大數據 御系统·而目對國家核心數據 供達尤其嚴格。數安法列明啟動數處項目時,必須 遵守法律印倫珪要求·制定全流程数服安全管理制 度\*世外,数法交易服务中介,有织数核应理服务 商·国家安全撤酬因酬法特需要獲取數據·以及境 外司法被推翻政境内數據時作出了規定。可見數字 「短適已經長入一個蔬要提升數據安全治理能力・メ 要保持數據高度流過及使用的時代。自2009年大 數據開始被提出,數據安全治理由2017年從寬鬆 · 转越最格管控·然後進入2021年筆者所指的數據

對於人灣亞來說,這又是來着什麼了內地 [十四百規劃網要] 拔出拓展數字灣區,並落實建 文為透淡障視人就發中心的政策目標。未來的數字 建論消支超影問依賴數據季報完通的接直情況,香 海和與門作為人將活內的核心候外城市,筆者認為 人名作為,在一层的全國演樂研究會中特別認出 有必要對對影樣跨球所動的不構通,導致影響民 生,當後企業創新發展進行深入研究。

封閉式防放的年代。

反動全球形列,以歐美為主導的數據安全規

#### 第日記會成為最內互聯調產業的最大發展降极,就 在二世美國政府以數要安全為由,是出對中國就 件面進行夏威将審查。在大國是杜架下,范濃維統 積極利用「國事副」起势,透過先行先就得後,愛 損金銀務處數應種種,私履保護,記息安全、監管協 展告部是陳手等機遇,而在大層百道幕鏈號成的數 操行是一,是建立安全跨場最佳的孵化也高,亦是 在安全與範的符號讓流過大前提下,探索理想模 式的好時機。

该信期間,以長居廣東省愛香港醫療資助的 港人難以巨港試斷為例,香港大學深圳醫院為這批 納志者提供診存服務,保成了港深層也病所首次跨 沒互通,為此為深素建立醫療数處能信號開及漂作 提供了良好的契線。類似港次軍主數據活通的例 了比比省是,除着人灣區的發展,入口流動愈落須 整,生企證,保險,教育,科在,醫療等易尝下都擁 行推山高效的數處污過線例。

数字 經濟正在快速推進之際, 漆與可望成為 注動跨過數據流通的約新餐口, 在完善人灣高的四 通《人派、物源, 主定、訊息定》之餘, 也可以半着 簡雜, 成為平認數據流通的特面。 軍品覺 紅杉資本中國專家合夥人、

11/2月平平國寺家古参入 阿里巴巴商學院特聘教授暨學術委員會委員













Federated Learning enables mobile phones to collaboratively learn a shared prediction model while keeping all the training data on device, decoupling the ability to do machine learning from the need to store the data in the cloud.



https://ai.googleblog.com/2017/04/federated-learning-collaborative.html





# Chatbots for Customer Services ... Most Common AI Application



### AI powered chatbots for self-service support workflows



of organizations are planning to deploy AI within the next three years. By 2030, a billion service tickets will be raised automatically by customer-owned bots. The trend of customers assigning their endless digital activities to their virtual personal assistants (VPAs), chatbots and other self-service tools will grow over the next 10 years

Note: She take a Commission behavior. Settyme Aura Without Instance Commission Auranne.



### Al Chatbot – Virtual Assistant (since mid 2018)





# You are using AI everyday



the

social dilemma

### Four waves of Artificial Intelligence Applications



Four waves of AI, according to Kai-Fu Lee



# AI is a game changing technology & will bring new order



# 世界人工智能雲端峰會揭幕 李彦宏:中國勢掀AI化

人工校会会会议 医麻痹 新知道 计图示器 上或抓服-多电料理应可算道-分辨施派 THE REPORT OF A ##人用書·拼形度[[em]]引直建建用 股寬(Stan Month),以其否定董事业兼备 建最利用量素完成的基础通信系统建一条 济于出示·财产中国立场和保证的利益性 社廳) 相關藥成者人類自能入層的經濟則 好好日来在约束大能使到1000年11日4日 大概兼任利益事業時候金額人の報告

平原创放5·人联队与 把就将被当 大学文指曰:但因性者能出于-被男相用 计算机量量量化,进展测量和稳定素光谱 其\$P\$10元6年建长港车行机通道时期一位 表示,当常我的人工型装饰常一颗封竹都 総約予由型の同一定部に対応確定局價目 第一日前月2010年代日前市路路路路 建作业公式第一条车来建立电路化来设置 統定的対象もしの

未建艺工化:1、人工製造県部沖湾 非·主体粉贴集;起为卫生建筑供信的经 建建筑-光以来封信河马车的有限(沙通星) 建建装装有:截留装饰的量10%差30%。 約歳付領事・原系に注意教養の工作に定 1236.增具、封持建备相过,自能影响车 (二)带来过行之前十年增加其自, 当注捕肉出,人工智能高人想要坐鲜

TRACK! NAAAAAAAAA###D-RomMAAAA

油:白河东市内-前部林台集市会 信息 特别人员的使首员副者,那是主义工程的 兼性问题, 记载, 记刻众日集, 生封州外袋, 員-位謂稱 确认专来建却把其一出礼能 take data and the second rate and add data 上海: 出加快帮助律师人, 密西上用人工

风影生和迷念-白皙就爱法告於中部 常常在 法课证法 原葉種様、素品に新生な工作問題(て適切) 金沙城,林川建大城城市和5 (第三百首席版):约许有笔址并不动乱;46 ※於京湖市建築・当時江間町研修新一行 清理道:《大類記由注系》書堂編輯(加) 人林道理自然和印刷面的低力能力。但是 美丽包稿-特别就好上海建建的起源 工業或詳述科、東点生工業新使用以注意 例如何都定该相同的你正,你可能能能能 ※ 令人知道过于我了去



101010-0010-0010-00

约42.811年位:约21一平通过近:

# Wave 3: Perception AI





# AI 金柱夏 ... one minutes = 1000 words



AI 人工智能

南韓電視台AI主播 24小時 on air 最近南韓 MBN 電視台推出首位人工智能 (AI) 主播,100% 還原令網友嘖嘖稱奇!

AI研發公司以MBN 電視台的著名女主播 金柱夏做藍本,仔細分析她的動作、聲音,研發 出AI版「金柱夏」,AI版「金柱夏」最快可以在 一分鐘內超高速生成一千字的報道。不少觀眾 見到片段後,都紛紛留言稱「超逼真」。推出AI 主播是希望當發生重大災難性事故時,可迅速 進行報道,最重要是可以24小時on air,對電 視台在人手、時間及成本方面,都有好大幫助。



Pay attention at 00:20 08:00

Pay attention at 02:30 03:20

https://www.youtube.com/watch?v=I82PxsKHxYc&t=18s

https://www.youtube.com/watch?v=5tvmMX8r\_OM&t=60s

# AI Trend – US vs China

# Wave 3: Perception Al



AI 人工智能





BAAN

北京智源人工智能研究院

Pay attention at 00:35 08:00

(GPT-3)

https://www.youtube.com/watch?v=PqbB07n uQ4

https://www.youtube.com/watch?v=mMI7rLETnQ0





### First Phase of Al





Second Phase of AI Empowered by Big Data





AI 人工智能



Symbolic AI, e.g. <mark>expert system</mark> Human knew how to approsimate knowledge with human rules

- Machine Learning involves the development and evaluation of algorithms that enable a computers to extract function from a dataset.
- Function (Inputs) = Output; function (5,3) = 15;

multiplication is the best choice.

machine learning algorithms can learn not only learn from the complex rules of their environment and also by being exposed repeatedly to more and more data

Source: https://medium.com/infinicog/current-state-of-artificial-intelligence-3d351a464425

# HONG KONG FINTECH WEEK 2018 HIGHLIGHTS Wave 2: Business AI 商業AI

Hidden Connection Weak & Micro Features



AI 人工智能

Markets Tech Media Success Video

୍ହ <u> </u>



This startup uses battery life to determine credit scores

by Hope King @lisahopeking () August 24, 2016; 4:00 PM ET

RISINESS

🖪 Recommend 72 🙆 🚹 🕑 🍈 🚥

"Online social and [mobile] behavior are predictive," Stewart told CNNMoney.

While the battery level doesn't matter too much in real time, the company looks at how that changes over a specific duration -- that can convey how consistent someone is and how much they plan ahead.

Lenddo puts the data points into a complex proprietary algorithm, which computes how likely someone will default on a loan. Lenders then decide the default rate they want to accept.

The trend of using smartphone data to make financial decisions is expanding in other categories, too. Lenddo is already helping companies like Globe and Indosat, two telecommunications networks in the Philipines and Indonesia, become lenders.

#### Zhaoli Meng, Dean of JD Fintech Research Institute at JD Finance



Dr. Zhaoli Meng is the dean of JD Fintech Research Institute at JD Finance. She is a graduate of the National University of Singapore and has over ten years of research experience in e-commerce, telecommunication, corporate management, and information systems.

Meng has numerous research reports published in China Daily, the Harvard Business Review, Peking University Business Review, and other notable publications. Her book Management Economics has been used as teaching material by many of the top MBA

programs.

ch Institute, Meng was chief economist and vice dean of Tencent Research Institute, IT Center of Excellence, research leader of Accenture Institute of High Performance, ng Economics Research Institute, and assistant professor at the Renmin University of





https://money.cnn.com/2016/08/24/technology/lenddo-smartphone-battery-loan/index.html

https://emerj.com/ai-sector-overviews/artificial-intelligence-applications-lending-loan-management/

### AI, Deep Learning, Image Recognition – All Explained

### Wave 3: Perception Al

Facial recognition for images is getting smarter

With Custom Classifiers, we train and deploy custom AI models for enterprise-specific labelling a



AI 人工智能



FACE DETECTION

The old approach: image recognition was difficult and less accurate



### Deep Learning is more accurate and can generalize concepts



### Product Categorization & Logo Detection can add enterprise value

FACE DETECTION



INPUT



PRODUCT DETECTION

LOGO DETECTION

Source: Adobe

# Geoffrey Hinton, "Godfather of Deep Learning"





## Al pioneer Geoff Hinton: "Deep learning is going to be able to do everything"

Thirty years ago, Hinton's belief in neural networks was contrarian. Now it's hard to find anyone who disagrees, he says.

by Karen Hao

November 3, 2020

# Why Deep Learning?



AI 人工智能

Hand engineered features are time consuming, brittle, and not scalable in practice

Can we learn the **underlying features** directly from data?

Low Level Features

Mid Level Features

**High Level Features** 

Wave 3: Perception Al







Lines & Edges

Eyes & Nose & Ears

Facial Structure


Deep Learning or "Neural network" models of AI process signals by sending them through a network of nodes analogous to neurons.



AI 人工智能





M. Mitchell Waldrop PNAS 2019;116:4:1074-1077



# Al is a Transformational Technology

"Immersion 沉浸式" deep mental involvement in something

AI 人工智能



Trends ... 100%





# MIT Schwarzman College of Computing

"...every graduate should encounter computer science and AI before graduation... regardless of their chosen profession."







理大在下學年亦會將「人工智能及數據分析」(AIDA)及「創新及創業」(IE)兩類課程加入大學核心課程,所有 本科生在4年內須在上述兩類科目,分別選修最少2個及1個學分,以達畢業要求。



■網上圖片顯示,要求上傳寵物照片後,支付寶 會根據鼻紋資訊建立專屬電子檔。 (洞上圖片)

 內地科網巨擘阿里巴巴(09988)同
 素支村寶,宣布開放其寵物鼻紋識別技術,並夥拍眾安保險(06060),中國大地保險,把該技術首次應用於寵物保險產品上。新產品「龐物險」是支付寶保險產品,現已可接愛維,夠兩大類龐物投保。 上、首次推出針對人類以外的斷療保險產品上。首次推出針對人類以外的斷療保險產品,現已可接愛維,夠兩大類龐物投保。
 職物照片後,支付寶會棍據寵物鼻紋資訊,為牠建立專屬的電子檔案,完成投保程序。目後申請理簡時,系統可透過鼻紋資料,簡便核實寵物身份。支付寶貴表示,相關動夠直類似人類指模,不會隨動物成長而變化。據悉該寵物鼻紋類別技術,目前識別成功率超過99%,解決以往在寵物保險理賠過程中,犂以識別龐物身位的問題,相比植入晶片等傳統方式,新技術更可免寵物受傷害。

#### 針對貓狗 保額達2萬人幣

支付寶現為3個月至10周歲的寵物 約・儲推出「寵物險」,完成投保者可報 鎖寵物門診醫療費用。寵物險保障額度 方面,分為3000元(人民幣,下同)。1 滴元,以及2萬元三個等級。以年繳保費 399元為例,一年累計最高可報鎖1萬元 醫療費用,單次理賠最高1200元,不限 疾病種類,亦可多次申請理赔。

支付寶引遞數據指出,中國在2019年的龐物 數量接近1億隻,相關消費規模高達2024億元; 但內地寵物保險覆蓋率仍低於1%,讓逓於美國的 25%,日本的7%。鼻紋技術有助推動寵物保險產 品在內地進一步普及之餘,支付寶保險財產險負責 人王方認為,該技術未來也有望應用於城市龐物 管理,尋找走失動物等用途。寵物保險市場潛力龐 大,市場上現有產品不多。今年4月獲香港保監局頒 發虛擬保險牌限的OneDegree,早前亦推出了「毛 價保」龐物保險,讓主人直接透過網站為寵物購買 醫療保險,每年保障額高達3萬港元。

# People are Fascinated on AI with Pictures & Images

### 「AI畫伯」網爆紅 大頭照秒變油畫

最近大家在 Facebook、Instagram,應該都看到不少人紛紛將 Profile Picture換成古典油畫風格。人人時光倒流四百多年,一夜間 化身文藝復興時期的文人雅士,原來全靠一個名為「AI畫伯」(名為 AI Gahaku)的日本網站。「AI畫伯」能夠分析用戶上載的相片,再透過人工 智能與不同風格的西洋畫結合。整個過程就好像手機相片後製程式般,只 需上載相片再點擊不同的風格,就可以由人工智能生成屬於你個人專屬 的西洋人像畫。

在進入網站後點選 TRY IT NOW,然後再選 Select a photo from your library,之後就能夠在10款不同風格中作出選擇。下載圖片後就可 以把它設定成fb或是IG大頭照,現在就來玩玩看! (網上圖片)





■超工尼研究的新AT模倣技训,宣稱能透過 網絡,自動轉換照片和影片內的人臉。

## 迪士尼AI換臉達高清級數

人工智能(Al)「易容」為近年熱門玩意,迪士 尼研究中心(Disney Research)及蘇黎世聯邦理工 學院(ETH Zürich)的研究團隊,近日開發出一項人 臉替換技術,能透過神經網絡自動轉換照片和影片 內的人臉,解像度更可達百萬像素高清級別,為觀 眾呈現逼真細節,有望於影視行業發揮巨大作用。 解像度過低是現有換臉技術的通病。研究團 隊對照片的人臉裁剪,再把人臉解像度統一調整至 1024 乗 1024 像素:經處理的圖像會傳送至通用編 碼器,把零散影像解碼,最後湊成一張完整人臉。

研究團隊強調,比起現時常見的靜態圖像,新 技術在動態影像的換臉效果更佳。從團隊釋出的實 驗短片可見,該技術遠較 Deepfake (深層偽造影 像)、DeepFaceLab的 SAE 模型等現有技術更逼真 自然。團隊指新技術能在影片重現昔日巨星風采,或 直接讓演員與特技替身換臉,省減後期製作工序。







# Fei-Fei Li



SEQUOIA CAPITAL PROFESSOR, CO-DIRECTOR OF THE STANFORD INSTITUTE FOR HUMAN-CENTERED ARTIFICIAL INTELLIGENCE (HAI) AND PROFESSOR, BY COURTESY, OF OPERATIONS, INFORMATION AND TECHNOLOGY AT THE GRADUATE SCHOOL OF BUSINESS

**Computer Science** 

## STANFORD VISION AND LEARNING LAB

The Stanford Vision and Learning Lab (SVL) at Stanford is directed by Professors Fei-Fei Li, Juan Carlos Niebles, Silvio Savarese and Jiajun Wu. We are tackling fundamental open problems in computer vision research and are intrigued by visual functionalities that give rise to semantically meaningful interpretations of the visual world. IM GENET

14,197,122 images, 21841 synsets indexed

Explore Download Challenges Publications Updates About

Not logged in. Login I Signup

**ImageNet** is an image database organized according to the **WordNet** hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. Currently we have an average of over five hundred images per node. We hope ImageNet will become a useful resource for researchers, educators, students and all of you who share our passion for pictures.

### **Emotion AI /Affective Computing**





https://www.startus-insights.com/innovators-guide/5-top-emerging-affective-computing-startups/













#### Emotion Detection

Detect and identify the learners' emotions and performance in real-time to understand where one struggles on particular subject



### Performance Prediction The Al-expert system able to analyze users' progress and forecast their performance



### Unique Algorithm Adjusts and rectify levels of teaching materials, narrowing learning differences



#### Patented Model

Cutting-Edge Ai technology incubated exclusive self-developed software with 3 patents granted and 29 patents pending



Squirrel AI Learning is the first pure-play AI-powered adaptive education provider in China.

We provide the personalized and high-quality K-12 afterschool tutoring at an affordable price.

We are addressing burning needs in education today.

- LACK of Personalized Attention in Traditional Classrooms
- UNEQUAL Distribution of Educational Opportunities

### About Adaptive Learning

Adaptive learning is an education technology that can respond to a student's interactions in real-time by automatically providing the student with individual support.

-Pearson X EdSurge "Decoding Adaptive" Report





### Now, AI Makes Online Courses Even Smarter



Joe McNamilrick Community of Al & Big Dune (read has achieved an examine more marked on frames

### Can AI Powered Education Close The Global Gender Gap?



Katte Elizabeth Commune co shark between defined building thematics (2010-b primers)





# When Seeing Is Not Believing



None of these people exist. These images were generated using deepfake technology. THISPERSONDOESNOTEXIST.COM

https://www.forbes.com/sites/robtoews/2020/05/25/ deepfakes-are-going-to-wreak-havoc-on-society-weare-not-prepared/#574294ae7494

# GAN generative adversarial network

### When Seeing Is Not Believing



The boom of deepfake technology is attributed to generative adversarial networks (GANs) introduced in 2014 by Ian Goodfellow.

#### GAN PROGRESS ON FACE GENERATION

Source: Goodfellow et al., 2014; Radford et al., 2016; Liu & Tuzel, 2016; Karras et al., 2018; Karras et al., 2019; Goodfellow, 2019; Karras et al., 2020; Al Index, 2021



2020

Source: <u>https://machinelearningmastery.com/impressive-applications-of-generative-adversarial-networks/</u> Source: <u>https://medium.com/infinicog/this-article-will-provide-a-brief-non-technical-explanation-of-what-generative-adversarial-41e47885f6a2</u>









AI 人工智能

Above: Synthetic images produced by StyleGAN, a GAN created by Nvidia researchers. Image Credit: Nvidia



Above: Synthetic images produced by StyleGAN, a GAN created by Nvidia researchers. Image Credit: Nvidia



Above: Synthetic images produced by StyleGAN, a GAN created by Nvidia researchers. Image Credit: Nvidia

# Trend -- As of now, AI is far from perfect Biases due to data and algorithm



AI 人工智能

080





Response: Racial and Gender bias in Amazon Rekognition — Commercial AI System for Analyzing Faces.



August 2018 Accuracy on Facial Analysis Pilot Parliaments Benchmark

98.7% 68.6% 100% 92.9%

amazon

DARKER

MALES



MALES



Amazon Rekognition Performance on Gender Classification



# Trend -- As of now, AI is not there yet

#### MIT Technology Review



https://www.technologyreview.com/topic/artificial-intelligence/

https://www.technologyreview.com/2020/04/27/1000658/google-medical-ai-accurate-lab-real-life-clinic-covid-diabetes-retina-disease/

https://www.technologyreview.com/2020/11/18/1012234/training-machine-learning-broken-real-world-heath-nlp-computer-

vision/?truid=086622dea1b423a3fb5237f65daeef9e&utm\_source=the\_download&utm\_medium=email&utm\_campaign=the\_download.unpaid.engagement&utm\_term=non-subs&utm\_content=11-18-2020&mc\_cid=658a1bfdeb&mc\_eid=0ee30778ed

#### COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK



#### <sub>据文</sub>:區結成醫生 牛侖倫理線 **再进中立大型十百倍理型中心遗乱**

今年我們中心籌辦一系列有關嶄新科技的倫理學研討,環繞兩 個主題:基因組醫學及人工智能的醫療應用。前者的技術和應用已 經頗成熟·相關的倫理課题和挑戰是清晰的;相比之下·人工智能在 醫療保健方面的發展,很多方面還在突破階段,技術在起飛中。它會 怎樣被應用於臨床診治,並不完全清晰。更大的特點在於人工智能 是具顛覆性的。可能催生全新的醫療模式。

一般醫療科技從研究開發到引入常規醫療,有既定的科技評估 程序,規範較容易;人工智能或不一樣:它不單是被動地融入既有 的療保健服務,更會從根本上改變範式,包括臨床診治的醫學邏輯。

以人工智能為主題的第一場網上碼 a conthmi? > 這立論不是全無爭議的 > 引 好好段前人傾該的那種交流恰恰托及 ! 討於4月中期利權行。兩位學者主講的 焦點都構確·波洲Monas= University 的 Robert Starrow 教授法同:人工智能經 深度學習大數據一口以根據演算法進行診 量) 经济的是一個具前期性的新研題。 治:在特定範囲,它的表現可以發展意料 醫生,在更小的時間可能誘過醫生的平均。」有高智慧,更形成感引時間 (affective) 2大道: 問題是、和賞人塾生一緒、古也會「対象對話、而是進化享舉聖に言、訪講話」 二起,病人觉情,能追究誰?一般的想 素表原態性,深度學習人類語言,可以問

有共同責任、問題只是如何分擔責任。 Spanow開的是一個哲學上更根本

的問題:雪人工若能被填予相當反該治難 **生的角色或複雜,而且它點向行眾度學育** 和设建自己约浦等法。内市以有其程度上 被視為一個「過德主题」呢?意思是。99 取對器空宴「自己」負責・不要察責其他 入 = Sparrow 的基本立論是不能+也不應 該一面發音和使用者小可發展於沒有人能 完全時館常振動|黒釉油草」(blac, bbx

||上昇・庶用人工智能診治・出了車、生産|||||放此地駅天、譲病大阪心||板線心車」。講||||量前注論性文章・覺得可以借來示定。| 而,推式設計員,可人利收集人工智能的 各洲心境出的防健是,要導入工智能(可 下人工智能在警察範疇的倫理課題,3位) 甸原機構,以及依賴人工智能的關件。詳一,以是機械人也可以是流動性式。成為病者,一作者之中,确位是涉入緊急的;加私傳一想,人工智能可以從根本上<del>成少</del>人為雜損, 前感性愛愛吃 答案並非目間的。 出錯責任 變得複雜 現代醫療中的醫患關係已經張小 空街灌坊人胡说心事。简庆恺的清澈 器座旗书》,通目也放在漫片器塑的人工

任,那不是自然不過的發展嗎?這個認識。

方向、與一股人設想未來人工智能會答問。

生做服恶性的声调工作、给醫生體出時間

动地位生者评論員Danie ligard(幕尼黑)

人工艺能被雇人日連省人化。它不但资

另一位 1 满具教Sarah Citan (集下)

和基言書教授(台灣)的原廣次提問:

到這方向的發展·講者未有作出知 治,但提出一话思考例了:在一個先導計 ·劃·醫师中心應用了人工智能向病人透露 理。這對學者在本文帶首的看法不詳而 诊断结果:「你想的是晚难愿遇+」在底来。 磁學·馬壺應消息(breaking sad news) 泉高難度動作·晉生未必は得很好·該有 quaiv)、例如它不只是靠活音询知進行。 智能、可以舒振的機械入下代码」 質素是 一 互性文解 否要有识證7團是非人性化?

在構想這篇文章時我遭到一篇有份 。專員支持用大律師和香港中文大學前校 造成的儒言·如果造成為趨勢·臨床的標 長·轉往至加坡委員的沈祖美教授。第二 法將把人工智能的服務標準沒統。在普通 一作名是Carboron Stewart法律學教授。 法地路 · 臨床標準一向是由原族團體自主。 hazard )… (Sydney University)。文章發表在《服算 制定的》法件上,應用人工智能的標準依 然可以遵循整要害士。 (很稀有·人工智能若能安病者的心,能 智能,但它的宏学法律和倫理性樂展示了) 图形、作谐物研媒和期释、病者及以信 更大的企图心:這是全面地討論人工智 能在設切聲來紀元的倫理和法律問題·包 括尊重強人主體、導私和保密、操守、成 方的責任領係躬得指於管理一在這個情況

作者認為,醫學會議中人工智能合務的復 人構想・人工智能感蕾・病人口多重交費 難問題當然有祥祭力地具證解決、但在當 的富贵派孩提供者可能會非依能管板環 前、施夕景好乐用晶的原则的方法云墩 推行更较近的思行保险計劃。例如可以通 派對人工智能的使用的权,設立资誉,把 合:在現階段一人工智能的相關倫這課題 ·现收用於距償受偿害的病者。而不需深究; 和状态或未好好地清晰界了…3位作者强 其中的過程是如何發生。

賠償了事 忽略檢討

然而也不能坐定,人工特益令各相籍

希學術文章應含成為此後開展討論的典 重者昔日常在腰營局主建體高高等 管理和核討醫療事故·封於不深研護器因 文章中有一點團帶可以呼燕一這 由·算邊提供方便賠償的機制 不免有些 Sparrow在我們研討會提出的1出位了實 保留(因為檢診醫療事故的目前)不實是該 究员任·更是英說法學習改進·防止可能 任許屬」的問題。人口智能統人醫療醫係 生、如果一旦个問答中因言、無異於完全 是會改變既有的責任服係。從好的方面 · 接受了 | 盖箱資算, 是無可奈何的事, 包 之由之 可能运長解忽了事所科技應用, · 治在採檢機念上也是讀德風險 (moral · 参考資料:Cameron Stewart Stephen K Y Wang, Joseph J Y Sund

Mapping ethico-legal principles fo the use of artificial intelligence if gastroenterology, 21 March 202 (intresi//enlinelibrary.wiley.com/doi/ pdf/10.1111/jde 155210 這利益衝突。微益。毋傷害和公平公下。 下,考虑責任該在哪裏是要熟悉目的。有

13人工智能的 入豐遼關係是

會改變既有的

責任關係。



AI ETHICS

MARK COECKELBERGH



AI 人工智能

From The MIT Press Essential Knowledge series

Al Ethics

By Mark Coeckelbergh

An accessible synthesis of ethical issues raised by artificial intelligence that moves beyond hype and nightmare scenarios to address concrete questions.



THE MIT PRESS ESSENTIAL KNOWLEDGE SERIES

### Stanford University Human-Centered Artificial Intelligence



CUHK Jockey Club AI for the Future Project 中大賽馬會智為未來計劃

# Trend – AI Power Houses are Pushing Back due to Ethics

·劉

The Futurist uncovers unique and inspiring topics that feed your curiosity. Curated articles on technology, science, and culture delivered to your inbox. Not for you? Unsubscribe anytime below.

/tech



## Over 1,000 AI experts condemn racist algorithms that claim to predict crime

Thousands of technologists and scholars from MIT, Harvard, and Google are speaking out against algorithms that attempt to predict crime based solely on a person's face, saying that publishing such studies reinforces pre-existing racial bias in the criminal justice system. The algorithms in question claim they can predict whether someone will commit a crime based solely on a picture of their face, with "80 percent accuracy" and "no racial bias," which experts in the field say is simply...not possible, **Vice** 



歐盟考慮未來3至5年內,暫時禁止在火車站,體 育館、購物中心等公眾場所使用人臉識別系統,以消 除民眾受到監視的疑慮。

### 保安項目獲豁免

網站Eurartiv取得歐盟長達18頁的人工智能(AI) 白皮書草案,歐盟建議在禁令期內,公營及私人機構 均不能在公眾場所使用人臉識別系統,歐盟將在此段 時間內研究有關科技帶來的衝擊,並制訂潛在風險管 理措施。不過,保安項目及研發計劃獲得豁免。

此外,白皮書建議 AI研發者及用戶均要承擔 責任,歐盟成員國亦應推出私隱及數碼權利的新規 例,以彌補現行政策的不足。新規例實施後,成員國 亦應任命相關機構進行監督,預計歐盟將在2月提供 最終版本。

德國計劃在134個火車站及14個機場設置人臉 識別系統,法國亦打算讓市民透過相關系統登入使 用政府網絡資源。不過,英國警方計劃使用實時人臉 識別系統進行監察,卻惹來政客及捍衛私隱組織抗 議,指摘現有系統不準確,措施侵犯私隱。

٦	10		BM停」	AL	人工智能
				834888 859899954 61195388 189728	
	1600員工促谷	AN DIGNI-	美警售朋	0007308	Residence of a second based on the second second second second second of the second second second second second second
	Charles N H2194 RAL 2017年1月11日 RALE 481 RALE 9 HER Laboration REF Transfer Fried RE 1 教育内皮采施数	Goog	ple:禁AI用於號	1164年1 (日日)- 祖 (日)-5月11-1	HARAN DI GA ANDAREN UN AN LARIS DA ANDREAM DI AN MARIT MARTANANAN KITA MARTINANANAN MARTINANAN
<b>亞馬遜辨人臉</b>	<ul> <li>NAME Information Contacts</li> <li>NAME INFORMATION AND CONTACTS</li> <li>NAME INFORMATION</li></ul>	Article and a second seco		ni Balitin Kanin C Kanin I Kanin I Citati	
(1)电影大学》,在11年1日中的日本中的日本中的日本中的日本中的日本中的日本中的日本中的日本中的日本中的日	K. 化热量过度在发展一系发展量、1000 等。因为1000年度本化中的管理和FASE和A 来用空气管理和Extensive 化可定用	1000000000	2010 (General Sec.)	100	保人權
		停售	長版	<b>會</b> 翻	別技術
					关端警察出售人脸流
		微軟總裁史	s密斯(Brad	Smith)	軟固四也加入行列。 明言,除非攀邦政府 保障人權,否則做軟
(		不會再向豐	隊提供這則	系統・	
					美國代理國家情報總 ETwitter的帖文·格
		當內蘭表示	- 既然微軟的	巨槌向漢	調警方提供人酸減剂
e È		超過8 別技術費給		e呼蘭科 當局用來	。 技臣頭不要把人臉讀 大規模監視人民 · 侵

## 三大因素推动人工智能发展:

## algorithmic innovation 算法, computation power 算力, and data 数据







2012

https://openai.com/blog/ai-and-compute/





## Fostering Big Data Applications in HK enterprises and Strengthening Collaboration of Digital Economy in the Greater Bay Area

- 公民數據科學家 Citizen Data Scientist
- 大數據潛力與數據經濟 Big data potentials and data economy
- 大數據治理概述 Overview of big data governance
- 數據治理原則 Principal of data governance
- 人工智能 Artificial Intelligence
- 機器學習 Machine Learning
- 深度學習 Deep Learning
- Jan 21 -- 商業建模 --- in the context of Cross Border Data
- March 4 -- 數據應用 -- Big Data Application Health & others
- June 24 -- The future of ecosystem of data economy -- Panel





https://hkinnovationnode.mit.edu/







