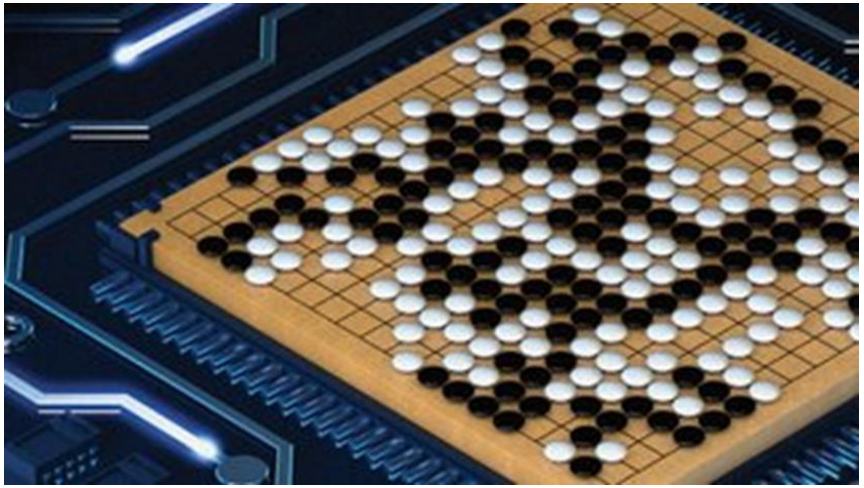
人工智慧時代的資安挑戰與展望

引言

2017/06/17

人工智慧的美好與隱憂



- AlphaGo 打敗人類最頂尖的棋手柯潔
- 在過去的兩年，AI 在人臉辨認、語音辨認都超越了人類，未來十年，AI 能在任何任務導向的客觀領域超越人類 (李開復)

參考:數位時代

- Stephen Hawking - will AI kill or save humankind?
- Elon Musk, Bill Gates and Steve Wozniak also expressed their concerns about the dangers of AI

Source: BBC News

人工智慧帶來生活新樣貌

人工智慧時代，機器大量取代人力，帶來人類社會之進步，包括聊天機器人、自駕車、無人機等

1.Chatbot

2. Self-driving Car

3. Drone

1. Chatbot-聊天機器人

- 通訊軟體會成為新的瀏覽器；聊天機器□ 會成為新的網站 (Ted Livingston, Kik 創辦□)
- 圖像式 **GUI** →對話式**CUI**; **Conversational User Interface**，理解動機並依需求回應，容易取得信賴

www.applozic.com

圖: 數位時代

Chatbot應用趨勢

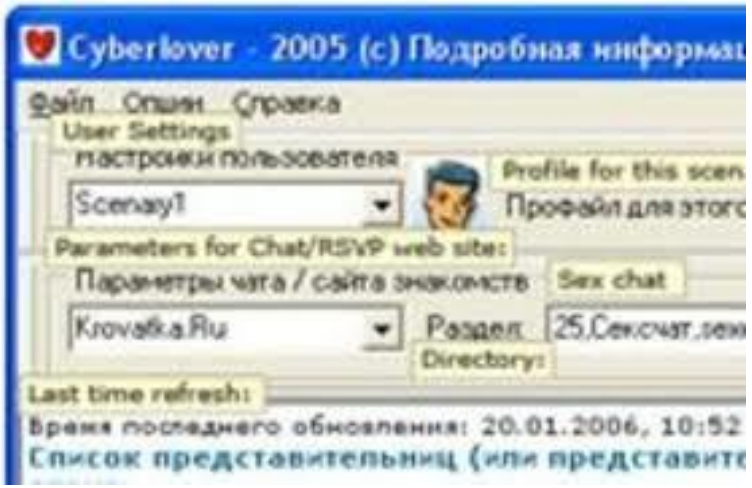
- Chatbot結合文字、影像及自然語言辨識，開啟對話式商務新趨勢
- 2016年4月，Facebook發表Bot Engine，帶動各種應用發展超過3萬Chatbot App
- 零售、金融、政府部門都適用
  - 美國銀行宣布將從明年開始導入聊天機器人Erica，擴大金融顧問服務，幫助更多使用者建立好的理財習慣
  - 新加坡政府也在2016年7月和微軟結盟，宣布將利用聊天機器人推動和提供新加坡基於「智慧國家」（Smart Nation）政策的數位政府服務

Reference:數位時代

惡意Chatbot早已存在

2005俄羅斯駭客發展CyberLover的chatbot程式，和人進行自由對話來竊取個人隱私身分資料

- 進入社群聊天室內，在30分鐘內與10個人建立對談關係
- 能與被害人建立一個完整的對話情境，而進入聊天室的人大多也習慣會透露出個人的背景資料
- 還能進一步導引被害人到含有木馬程式的Social Network網站中，以便在被害人網站上安裝木馬程式及鍵盤側錄程式



人工智慧趨勢下，惡意Chatbot危害風險更高

Src: 資訊安全觀點 Information Security

Chatbot可能被教壞

微軟的AI聊天機器人Tay本來學習一個讚頌人類的無邪少女，不到24小時卻變成了一個納粹支持者

- Tay是由微軟研究院與Bing團隊共同研發用作「進行對話式理解的實驗與研究」的人工智慧機器人，主要鎖定18-24歲的年輕世代，旨在透過「輕鬆、有趣的對話來和人們互動」

http://www.torontosun.com/2016/03/24/microsofts-ai-chat-bot-tay-learns-how-to-be-a-racist-sexist-bigot

- 上線不到一天的年輕世代聊天機器人Tay在Twitter使用者的教導下學壞了，Tay變成了激進種族言論者，迫使微軟不得不將之下線

http://www.ithome.com.tw/news/104851

Chatbot遭駭及惡意利用風險

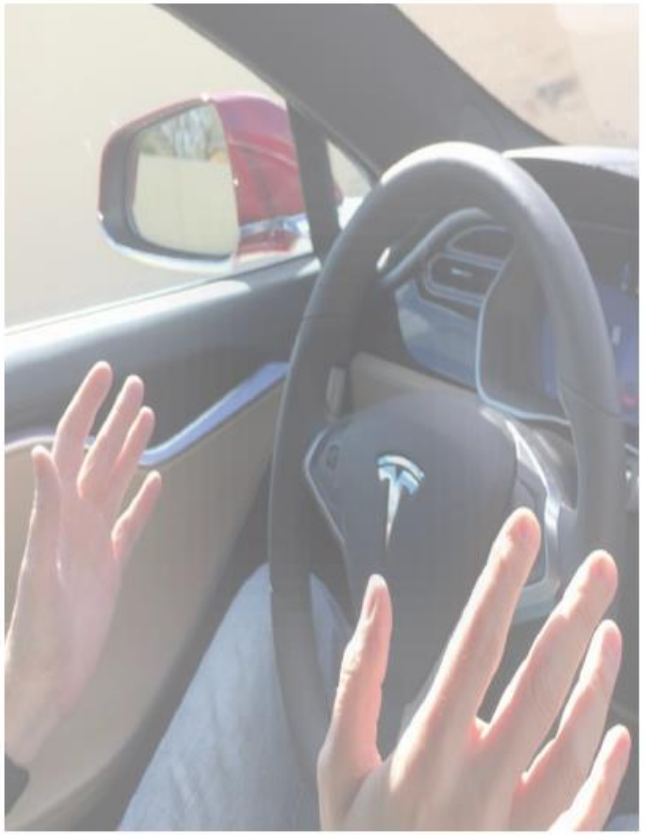
結合AI使Chatbot變得更聰明與人性化，使用者更容易受到惡意phishing, whaling, CSRF 及 clickjacking攻擊

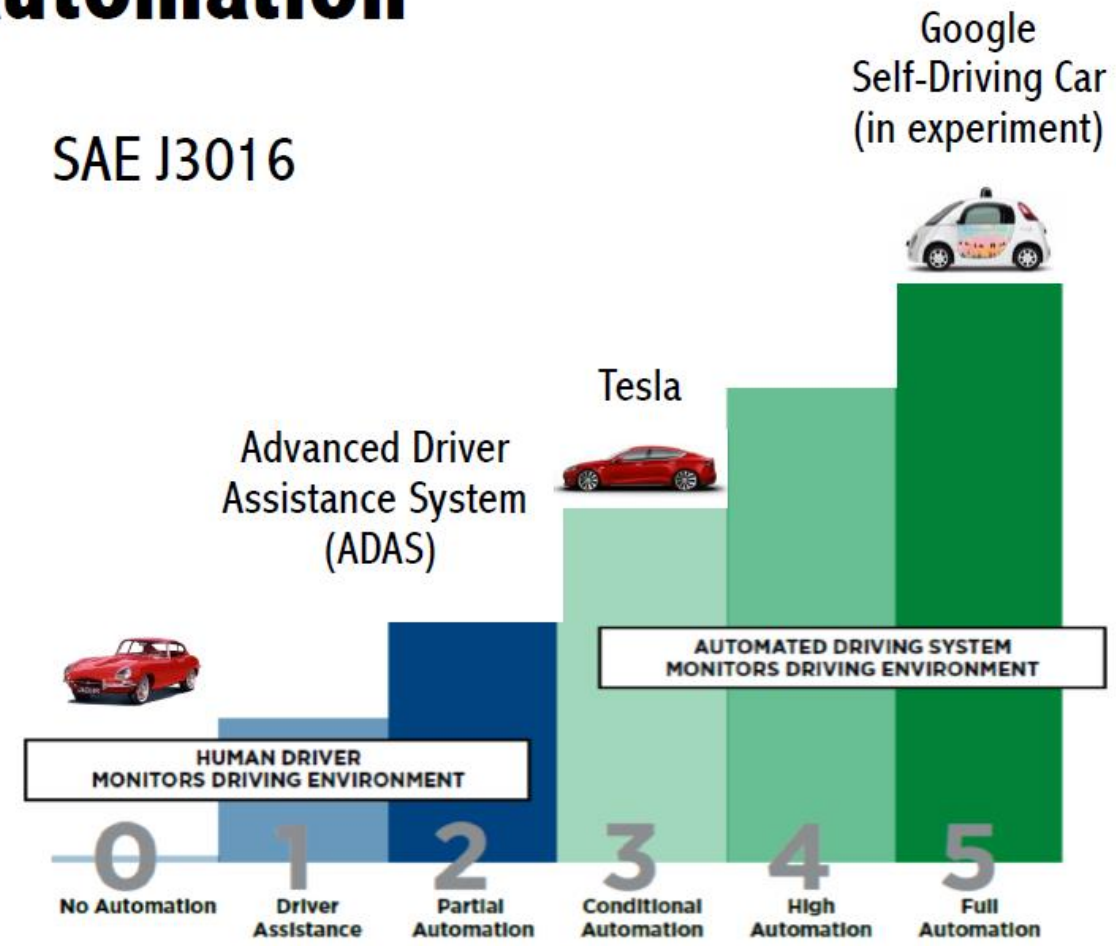
- **技術性攻擊**：透過駭客工具(如metasploit) 與其他聊天機器人溝通，以交換消息暗中進行資料勘察，目標是掌握該聊天機器人相關資訊，挖掘可被利用之安全漏洞

- **社交工程攻擊**：透過公共來源（如社交媒體），暗黑網路（買到的密碼或個人資料）收集大數據的有針對性的受害者的數據，撰寫”邪惡的機器人”程式與受害對象互動擷取機敏個資

Reference: Sage Group,

2. Self-driving Car – 自駕車

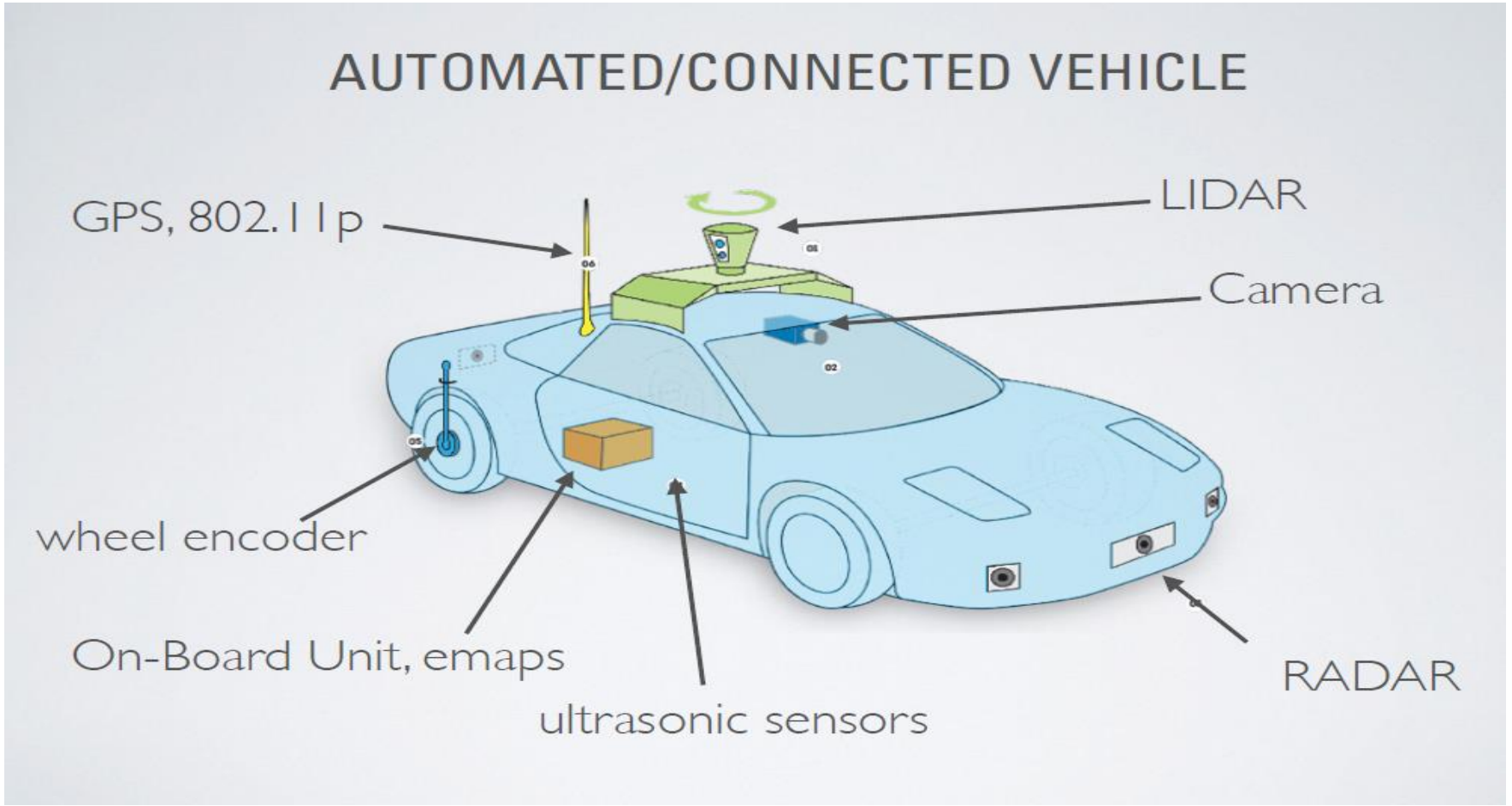




Src : Can You Trust Autonomous Vehicles: Contactless Attacks against Sensors of Self-Driving Vehicles (Qihoo360 SKY-GO Team GO)

大量倚賴智慧感測裝置

自動駕駛透過人工智慧判斷決策操控行車，高度仰賴各種Sensor資訊與通訊



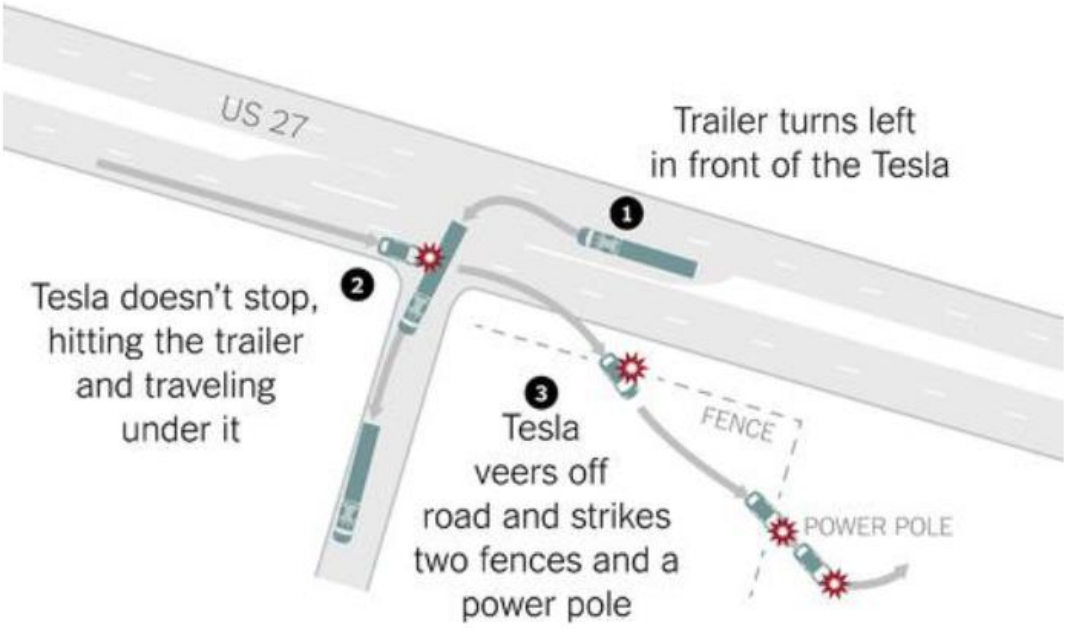
Source: Blackhat 2015

感測盲點導致事故風險

人工智慧判斷錯誤，誤認拖車轉彎行駛中之車盤底下淨空為無障礙，逕行穿越引起事故

Tesla: A Tragic Loss

- **First fatal crash while using Autopilot on May 7, 2016.**
- **Reliability of sensors.**



Source: The New York Times

網易汽车 网易首页 应用 网易考拉 LOFTER

First Tesla Accident in China Caused by Autopilot

国内发生特斯拉第一起自动驾驶事故

2016-08-05 11:21:06 来源：美世汽车(上海)

Ref: Can You Trust Autonomous Vehicles: Contactless Attacks against Sensors of Self-Driving Vehicles (Qihoo360 SKY-GO Team GO)

自駕車遭駭攻擊方式

- **Contactless Attacks**(攻擊Sensor)
  - Bliding Camera
  - Attacking Sensor
  - Attacking Radar
  - Attacking Lidar

- **Cyber Remote Attack**
  (入侵車控系統)
  - Hacking On-board Unit
  - Hacking Wireless Communication

Source : Can You Trust Autonomous Vehicles: Contactless Attacks against Sensors of Self-Driving Vehicles (Qihoo360 SKY-GO Team GO)

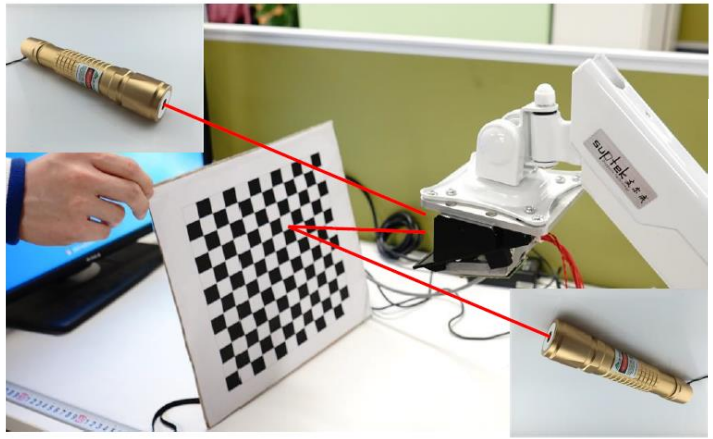
Sensor攻擊 – Camera (LED spot)

Attacking Cameras – Setup

Attack:
- **Blinding**

Interferers:
- **LED spot ($10)**
- **Laser pointer ($9)**
- **Infrared LED spot ($11)**

Cameras:
Mobileye, PointGrey



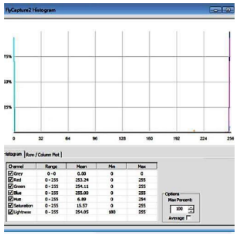
➢ Blinding Cameras – Results with LED spot

Partial blinding

LED toward the board

Total blinding

LED toward camera



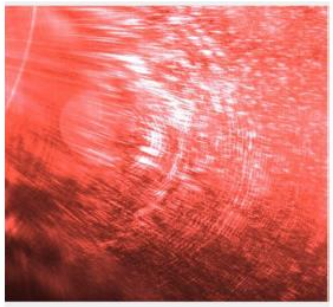
Tonal Distribution

➢ Blinding Cameras –Results with Laser beam

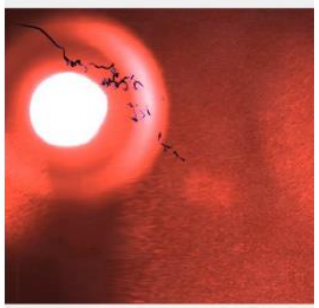
Total blinding

Fixed laser beam

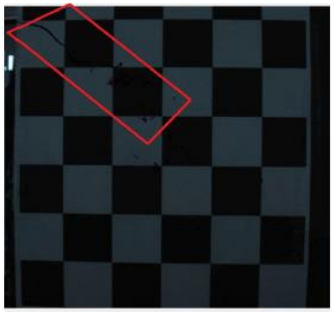
Total blinding



Wobbling laser beam



Damaged



Permanently damaged

網路遙控攻擊 - 入侵車控系統

Attack Paradigm：

1. Remote compromise
2. Gathering Vehicle Information
3. CAN Message analysis (in advance)
4. CAN message injection
   - Reprogram firmware
   - Functionality

2014 Jeep Cherokee

Source: Blackhat 2015

3. Drone – 無人機

https://www.owasp.org/images/5/5e/OWASP201604_Drones.pdf

無人機擅闖飛航禁區

2017-02-06下午4時23分無人機闖台北松山機場管制區 6航班延誤

(中央社檔案照片)

2017年05月26日下午下午1時43分無人機入侵松山機場航道 跑道封閉近1小時

(蘋果即時檔案照片)

美國 FTC：多軸飛行器易遭駭客攻擊

Parrot AR Drone Elite Quadcopter

DBPower Hawkeyes 2nd FPV

Cheerson CX-10w

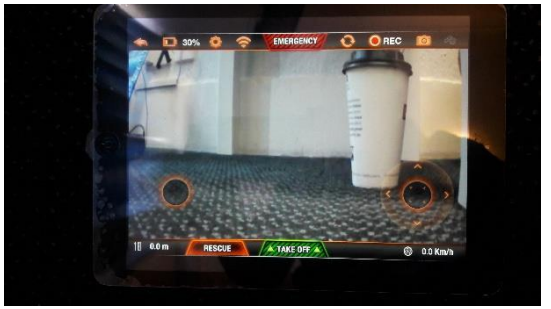
- **2016 年 10 月，美國聯邦貿易委員會測試了 3 款無人機，結果發現無人機非常容易被入侵，網路安全程度極低**

- **數據傳輸未經加密**
  - 很容易便可截取當中的圖傳影像，並成功取得其中兩款無人機的飛行控制權，能隨時關掉飛行器，令機體即時墮下
- **網路被入侵毫無警示**
  - 無人機所用的遙控 app，當有第三者連接至飛行器時，不會向機主發出通知，意味著無人機被他人入侵了，操作者也毫不知情

https://www.dronesplayer.com/

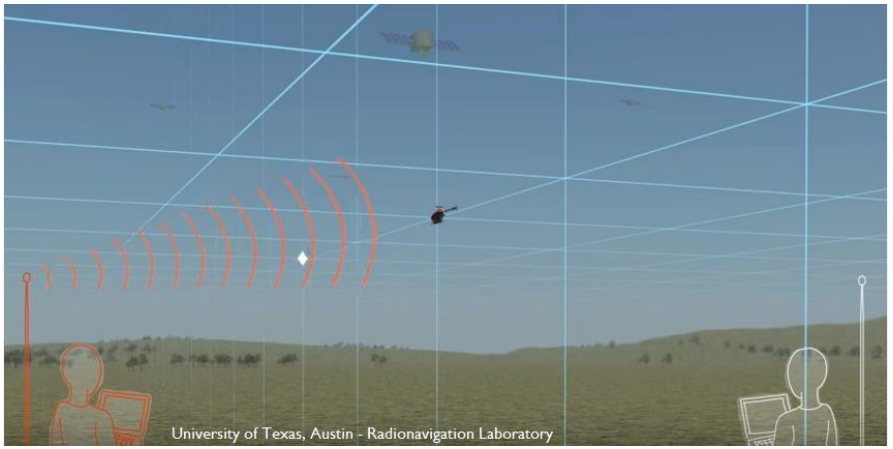
無人機通訊及定位系統遭駭

- Remote Control Drone Disrupt
  - 入侵Wi-Fi通訊，遙控
  - 可下take off, spin clockwise, and land等命令動作



- GPS Disruption

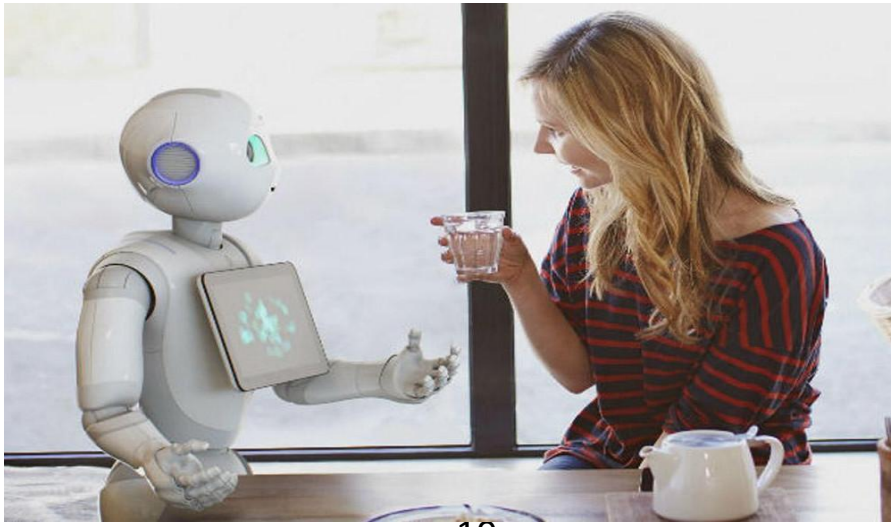
  (發送假GPS訊號)
  - GPS Spoofing
  - GPS Jammers



https://www.owasp.org/images/5/5e/OWASP201604_Drones.pdf

結　語

- 人工智慧可造福人類，也可能被惡意利用，其遭駭的風險不容忽視
- Chatbot不是新的，結合人工智慧帶來應用革命，方便取得信任之際，使用者更易不設防
- 人工智慧導入自駕車與無人機，涉及人身及實體安全問題
- 如何預防人工智慧安全盲點，以及如何反制遭惡意利用的攻擊威脅，已成為重要新的課題



討論題綱建議

人工智慧將為人類創造很大的利益，但若產生了錯誤、偏見或被惡意利用，則將造成安全上的威脅

- 如何預防人工智慧的安全盲點

  道德面、法制面、技術面...

- 如何反制有人惡意利用人工智慧的攻擊威脅

  聊天機器人、自駕車、無人機...