Kubernetes Summit '21

可觀察性對於金融業內的AI科學家帶來的好處與挑戰

宋政隆 (clsung@)
中國信託商業銀行

About.Me/clsung

- 中國信託數據暨科技研發處 2018 ~
  - 人工智慧科技研發
  - 數據治理規範管理
- Product Development, HTC Healthcare ~ 2018
  - Cloud Service Infrastructure
  - Mobile App Development
  - AI Platform (DeepQ)
- Open Source contributor (clsung@)
  - https://github.com/clsung
- LINE API Expert

Certified Kubernetes ...

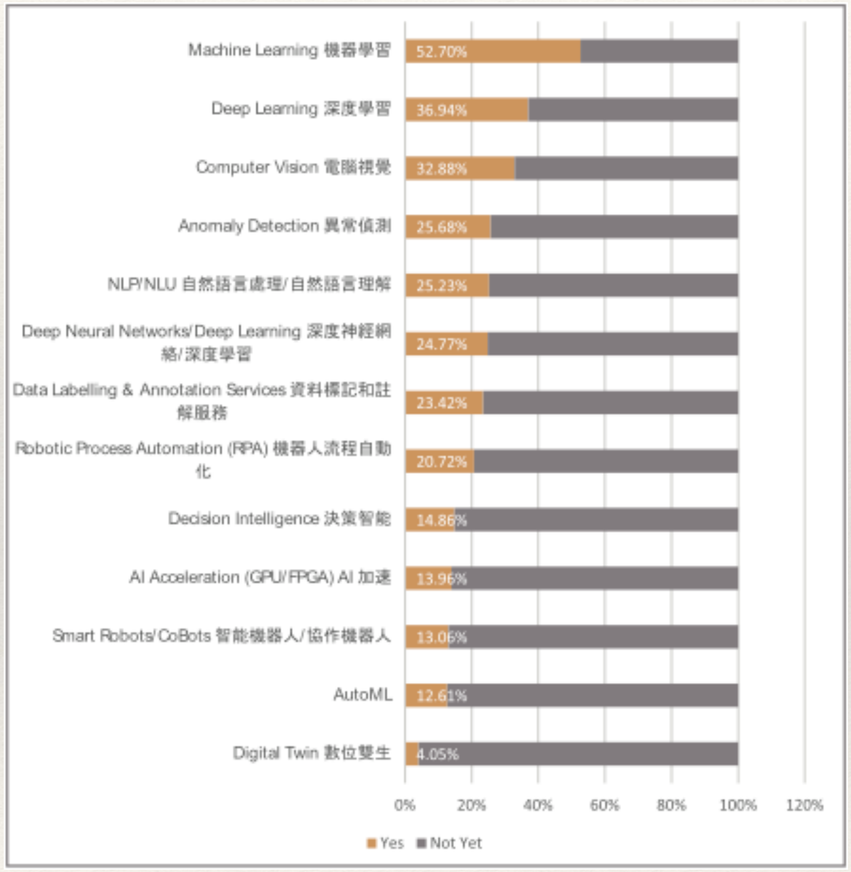
Agenda

- 中國信託人工智慧發展現況
  - 中信的 AI 應用在哪裡
  - 為什麼要選擇微服務
- 由 AI RD 的角度來看 Kubernetes 微服務
  - 選擇什麼微服務?
  - DevOps 與 K8S
- 可觀察性（Observability）
  - 如何協助我們在服務突然出現異常時，儘速找到問題?
  - 在使用託管或是受限制的服務中的技術挑戰

Kubernetes Summit 2020 觀眾輪廓

職階	職務		產業	
高階主管	工程師	架構師	資訊軟體與服務	資訊安全業
6.4%	41.8%	2.7%	29.8%	5.3%
中階主管	專案經理	資料分析師	系統整合服務業	交通運輸
8.7%	7.7%	2.3%	16.8%	2.0%
基層主管	維運工程師	DevOps 工程師	金融	電信與網際網路
20.3%	6.2%	1.6%	15.8%	1.8%
一般員工	技術主管	其他	製造業	其他
64.6%	5.7%	32%	6.3%	22.3%

人工智慧@台灣



- 各種技術都會交叉被使用以建立企業想要的AI應用架構
- 最常採用機器學習或深度學習的技術
- 與電腦視覺或自然語言處理結合應用時企業可採用不同形式的預測分析和機器人流程自動化來提升工作流程和業務效率

人工智慧@中信

自然語言處理

AML負面新聞判讀

處理 AML (反洗錢) 負面新聞，同樣的新聞篩檢費工費時，透過 AI 演算，不只能夠聚類單一事件的所有相關新聞，更能擷取文章重要資訊，生成新聞摘要，幫助業務單位更認識客戶，和輔助洗錢防制 AML 工作。

履歷智篩

只要按下按鈕，系統就會利用大數據分析，將大約1500份的履歷按照關鍵字做第一層篩選。過去需要花上三週的時間，現在只需要三天，就能挑出排名前百分之20的應徵者進入下一關。

刷卡消費體驗旅程

利用了自然語言處理技術，結合文字與時間序的特徵值，甚至，利用網路爬蟲技術蒐集外部資料，找出各種民眾慣用語...後續，中國信託將系統既有資料與網路爬蟲蒐集的慣用語比對，歸類出大眾容易理解的商店別名加註到刷卡消費通知上

電腦視覺

反詐騙智能ATM

臉部被安全帽或口罩遮擋、客戶轉帳時持續使用手機通話，將進行貼心提醒，預防詐騙、提升帳戶交易安全性；「智能影像分析」則透過智能判斷客戶個人化特徵如年齡、性別等，提供客 製化的互動訊息。

人臉辨識 (NIST 認證)

中國信託的微笑打卡功能，則可支援多人辨識，員工以活體人臉辨識取代員工門禁卡，並以「微笑」的表情代表確認，系統偵測到微笑就可完成打卡；另外還可辨識員工的穿衣風格，包括正式服裝、商務休閒、休閒服裝等類別。

AI OCR

自行研發的印刷體的文字辨識核心、手寫英數的AI辨識核心、文印鑑辨識技術，通通導入支票辨識上；...。此外，為了持續優化辨識正確率，中信更導入AI反饋機制，內部自己發展出標記功能，來改善標記效率...。

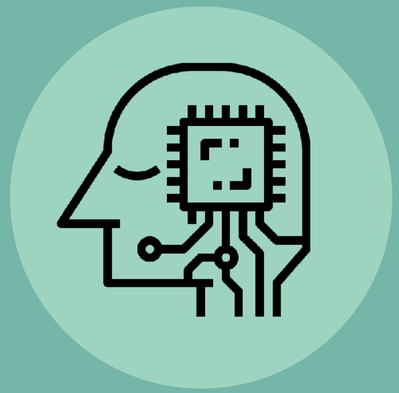
Ref: iThome, 各家新聞稿

人工智慧@中信

自然語言處理

閱讀理解

- 語意理解
- 文本解析



KYC

- 知識圖譜
- 主題模型

電腦視覺

文件辨識

- 印鑑比對
- 文字辨識

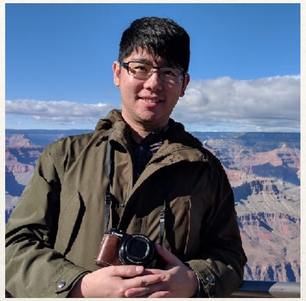
身份驗證

- 人臉辨識
- 物件辨識

人工智慧@中信

自問自答:
命名實體識別應用於精準醫療服務
- 以智能理賠為例

PyCon.TW 2021



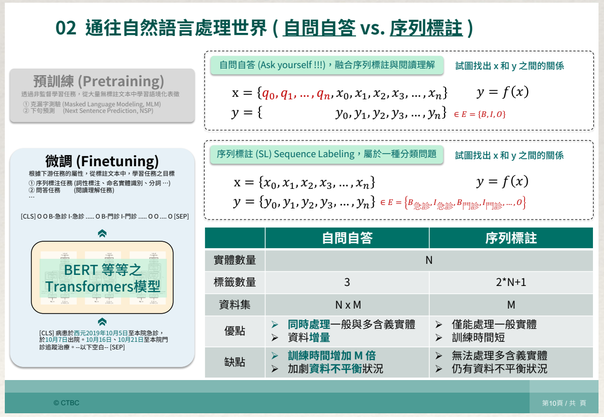
運用非監督式學習技法打造風險警示系統
以國際貿易作業場景為例

MOPCON 2021

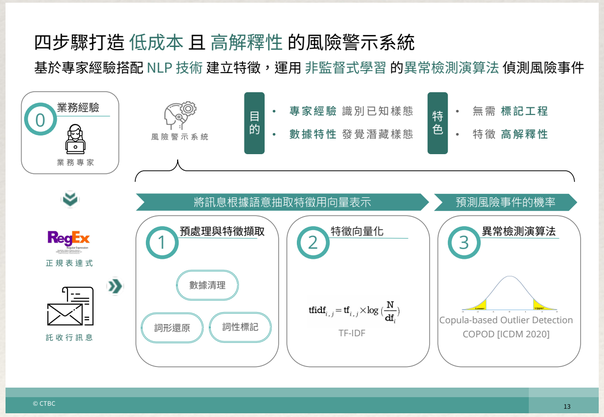
求改變或是求生存
中國信託自建 AI 平台的研發之路

Cloud Edge Summit '21

核心

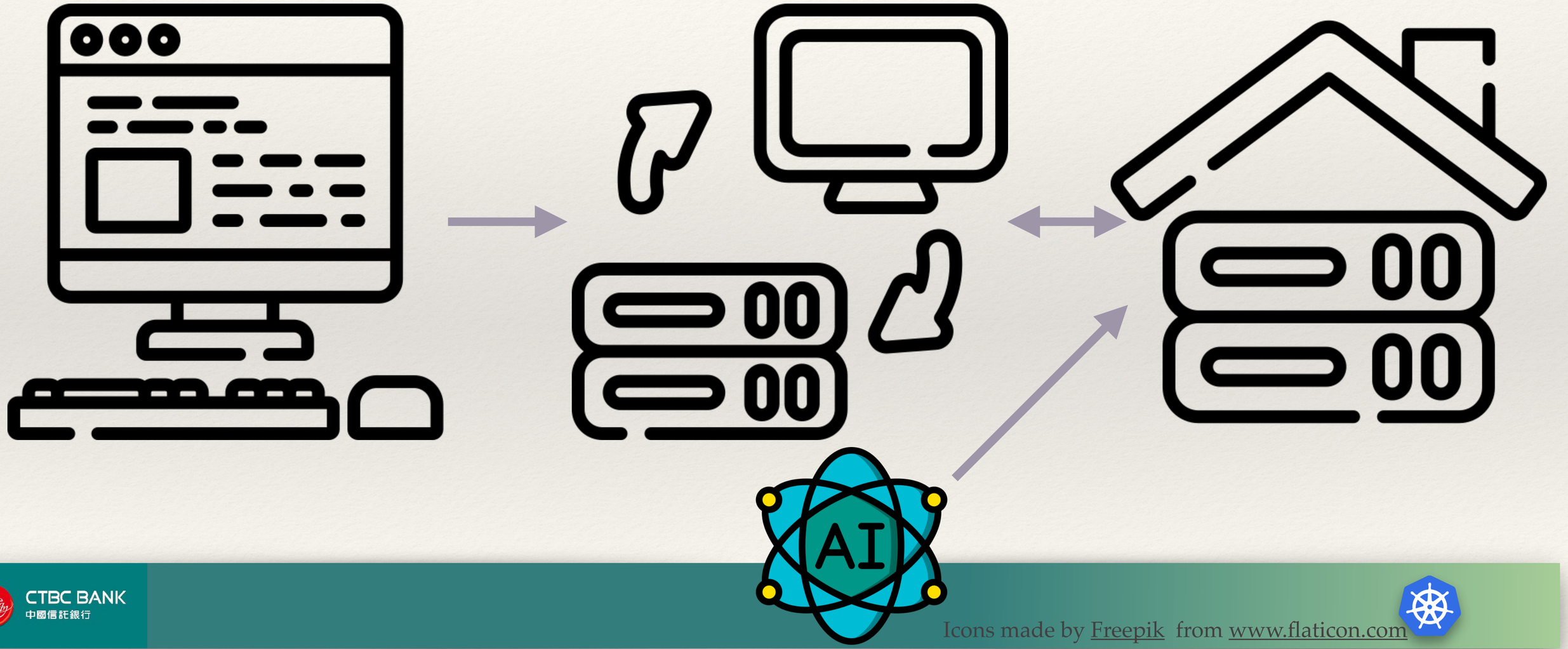


場景



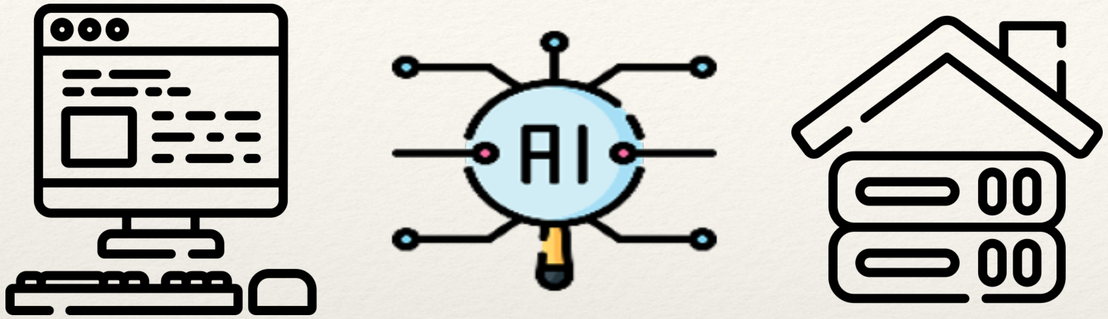
平台

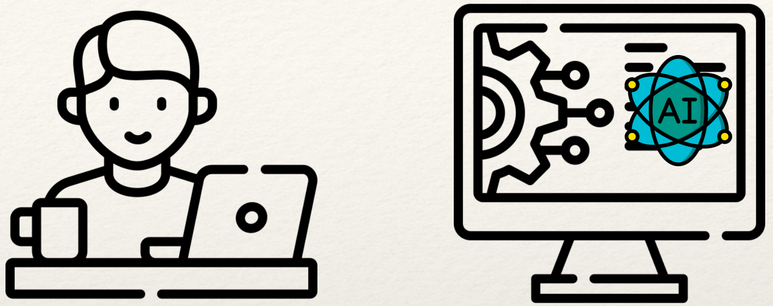
人工智慧@中信 專案一

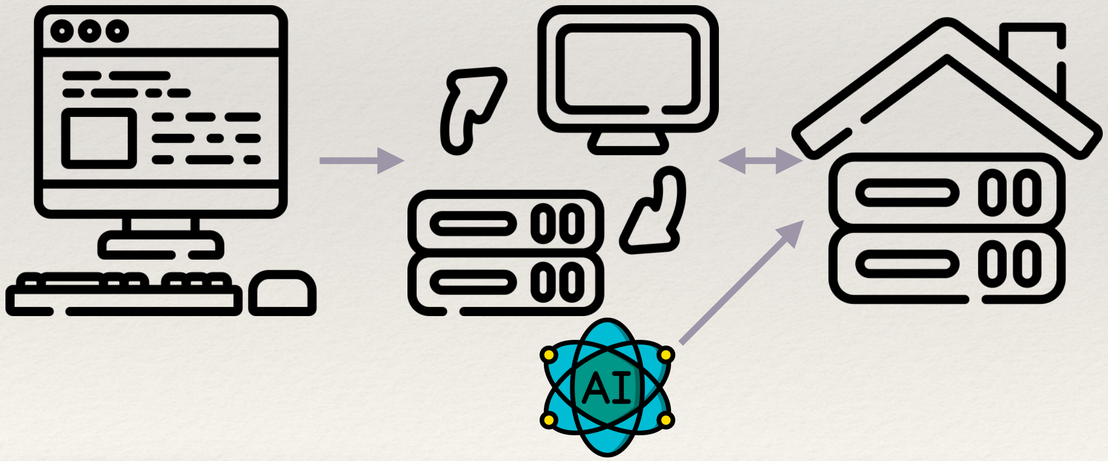


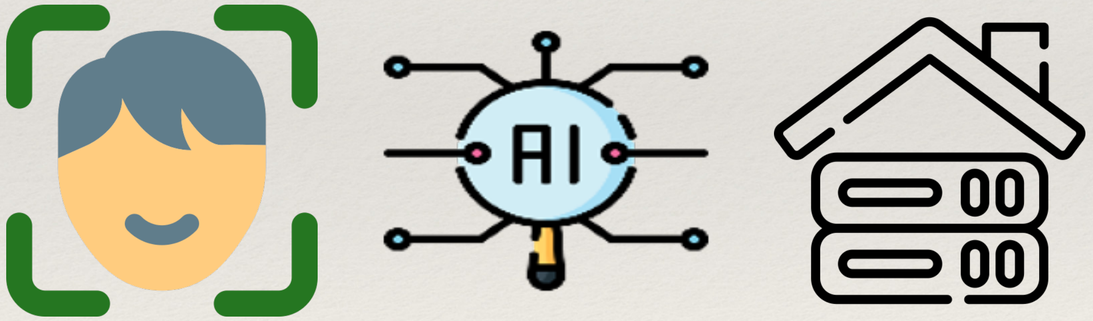
Icons made by Freepik from www.flaticon.com

人工智慧@中信 專案甲乙丙丁









隱私

個敏資的保護與使用

人工智慧@中信

AI 核心自建評估

共用

多項業務需求相近，一次開發、多次使用

彈性

業務需求彈性客製，快速迭代更新

創新

可進行多項實驗，支持創新業務

業務應用構面

風險控管

流程優化

行銷經營

AI 核心 + 人機協作 + 實驗場域

【AIGO】 全行共用平台

NLP Kernel

語意理解

斷詞引擎

文本解析

詞性分析

CV Kernel

人機協作介面

API

實驗場域

機器學習

標記工程

資料採集

模型訓練&優化

AI應用至各業務需求，單一服務平台

落地場景

法人金融

個人金融

行政中後台

金控子公司

人力資源稀缺

老闆說我們有千軍萬馬

千軍萬馬日理萬機

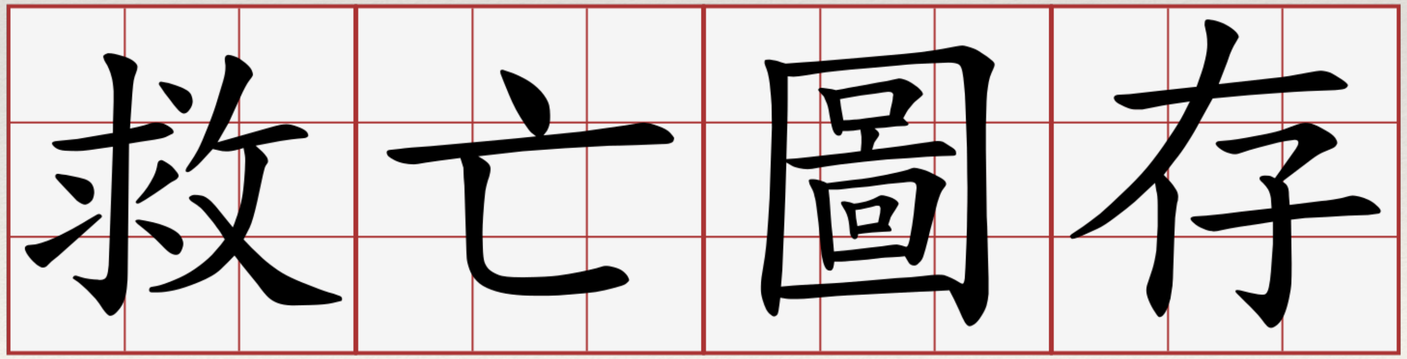
被打爆啦！

下個案子也要請多幫忙

UAT/SIT 都過了，但上線驗證失敗

面臨到的問題

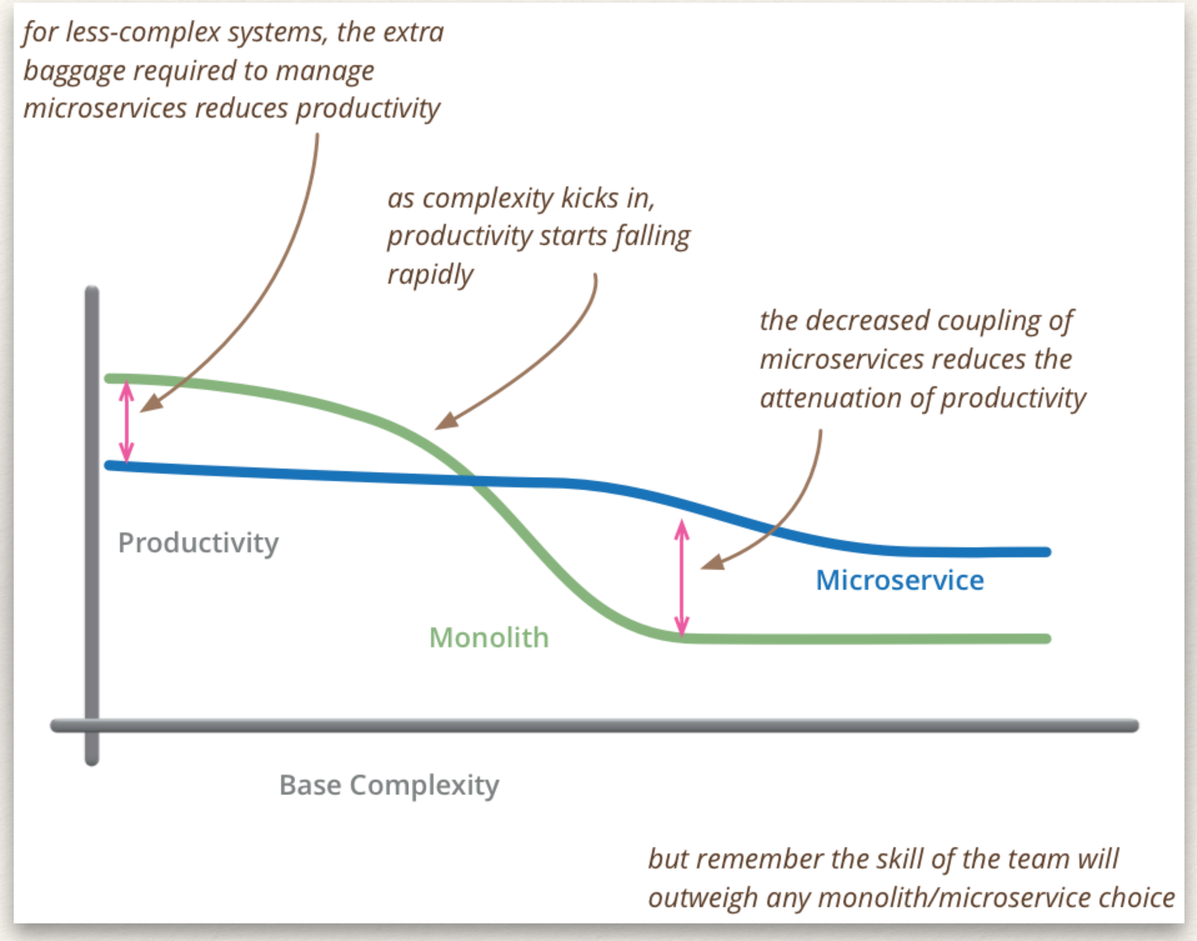
Q: 人力固定的情況下，做了**n**個核心，整併到**M**個不同的系統 (**M>n**)，
開發不難、部署有點難，版本更新好麻煩，出問題時 debug 好困難



為什麼要選擇微服務?

- So my primary guideline would be **don't** even consider microservices **unless** you have a system that's **too complex** to manage as a monolith. **The majority of software systems should be built as a single monolithic application.** Do pay attention to good modularity within that monolith, but don't try to separate it into separate services.

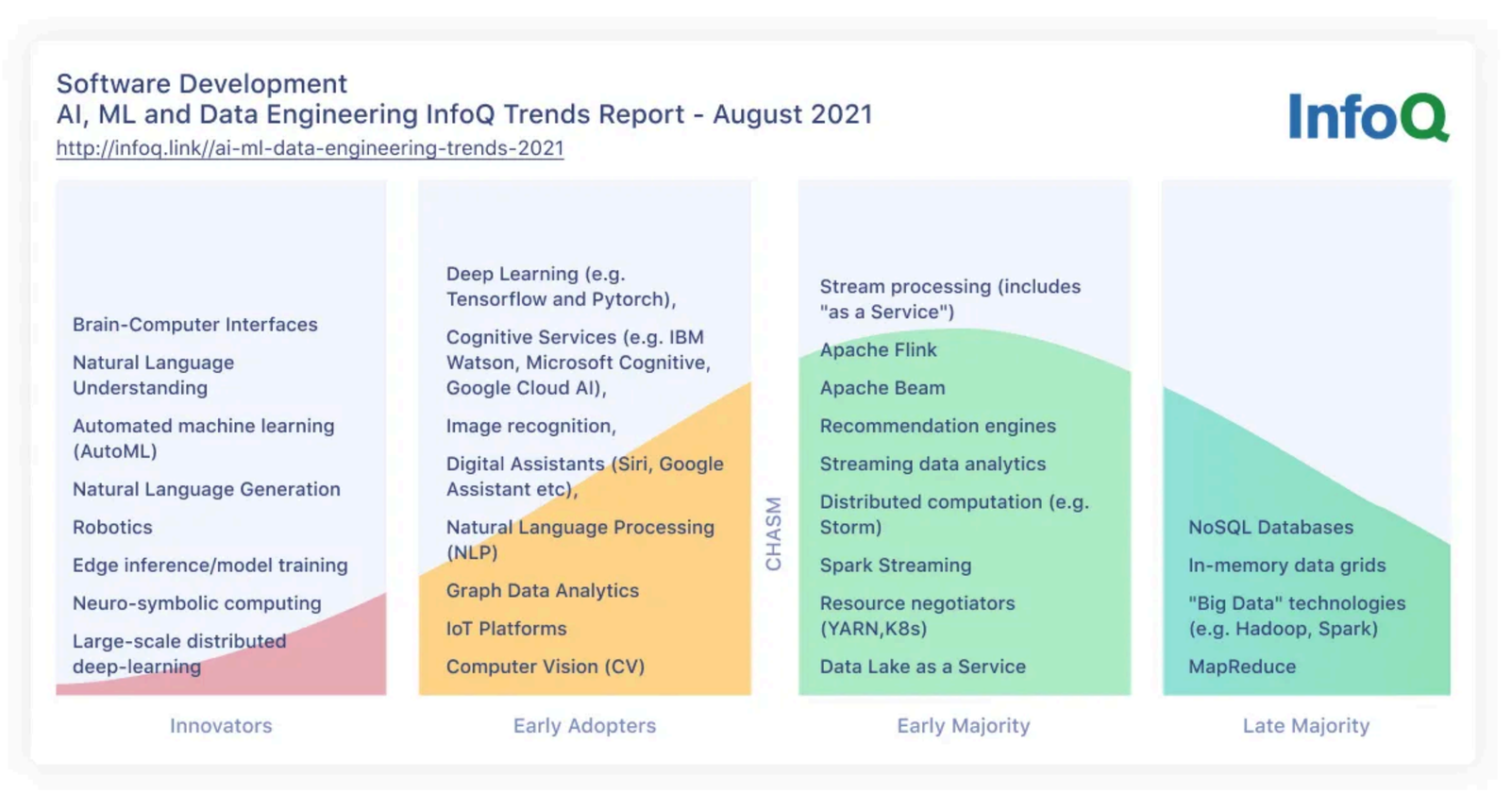
  — *Martin Fowler*



Ref: https://martinfowler.com/bliki/MicroservicePremium.html

為什麼要選擇微服務？ (now)

- MLOps and Data ops allow easy training and retraining of algorithms
  - KubeFlow
  - K3s, KubeEdge (edge)
- AutoML allows for automating part of the ML life cycle
  - MLflow

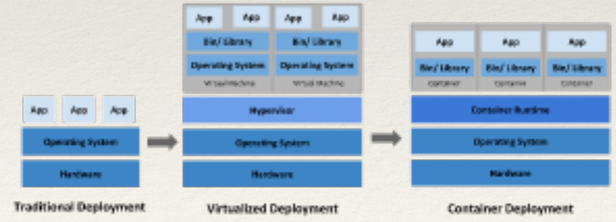


Ref: https://www.infoq.com/articles/ai-ml-data-engineering-trends-2021

為什麼要選擇微服務?

- DevOps / Sr RD 視角[*1]:
  - 缺乏可擴展性
  - 缺乏可重用性
  - 缺乏可監控與可觀察性
- Data Scientist /Jr (and some Sr) RD 視角:
  - 在我這裡跑是好的

謎之聲：不然我們來跑敏捷好了

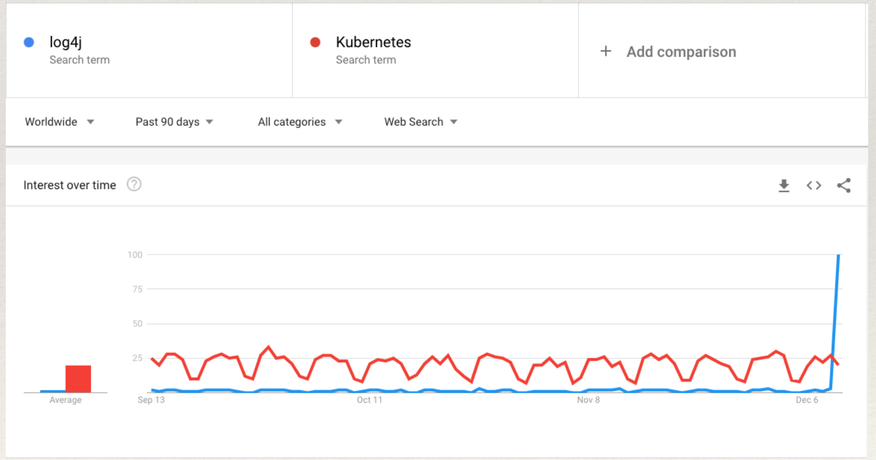


*1: 求改變或是求生存：中國信託自建平台的研發之路 - 彭士爵, Cloud Edge Summit 2021

為什麼要選擇K8S?

- IT /DevOps 視角:
  - 高可用性 (HA)
  - 自動擴展 (Auto Scaling)
  - 藍綠部署、金絲雀部署、滾動部署 ....
  - 因 CNCF，可監控與可觀察性有，且彈性大
- Data Scientist / RD 視角:
  - 著重在容器 (container image)的交付
  - 所有 AI/ML 核心，以 API 服務
  - 我不用再寫 Java 了....

謎之聲：微服務學K8S換工作比較容易



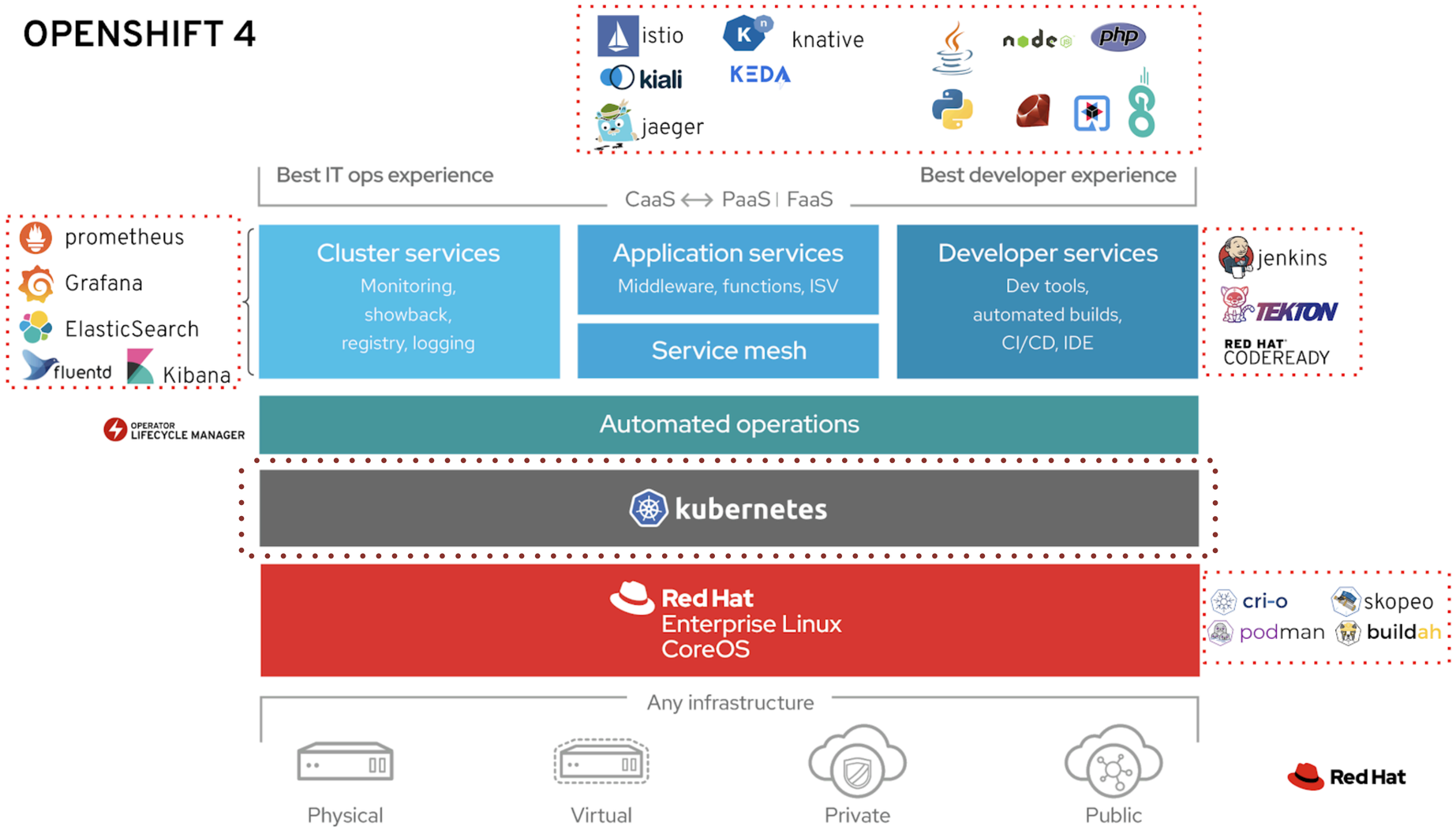
選擇什麼微服務?

kubernetes

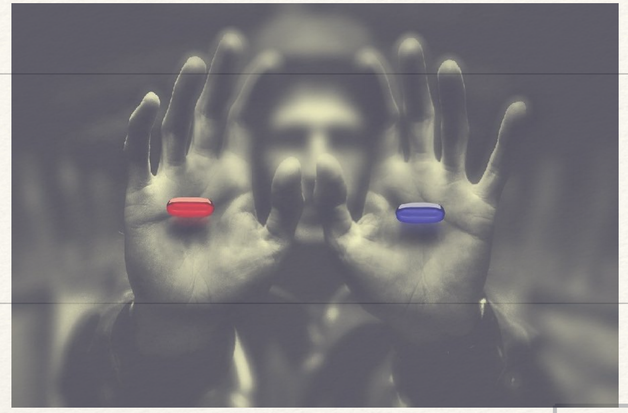
Red Hat
OpenShift Container Platform

kubectl commands	oc commands
kubectl get pods	oc get pods
kubectl get namespaces	oc get namespaces
kubectl create -f deployment.yaml	oc create -f deployment.yaml

OpenShift Container Platform



https://cloud.redhat.com/blog/enterprise-kubernetes-with-openshift-part-one



選擇什麼微服務?

OpenShift Container Platform

Pros

- 原廠 (RedHat) 的支援，24/7 小時服務不間斷
- 許多元件/Web UI已安裝好，可直接管理部署
- Security Polices 較嚴格，預設安全性高
- 穩定、可靠

Cons

- 需注意應用程式被平台綁架
- Security Polices 較嚴格
- 能安裝的套件限制多 (helm / operators)

Kubernetes

Pros

- 社群的支援，24/7 都可以上社群求救
- 支援各種作業系統，更新版本功能容易
- 部署原生應用程式容易
- 開發彈性高，不爽可以自己來

Cons

- 什麼都要自己來
- 安全性容易被忽略，如 0day (log4j 表示….)
- 相關工具需自己整合，如 nginx-ingress、prometheus…

選擇微服務的角度，是 DevOps 還是....?

謎之聲：小孩子才做選擇，我全都要

DevOps 與他的快樂夥伴

DevOps

DevOps是一種重視「軟體開發人員」和「IT運維技術人員」之間溝通合作的文化、運動或慣例。

ModelOps

主要關注各種可操作的人工智能和決策模型的治理和生命週期管理，包括機器學習，知識圖，規則，優化，語言和基於代理的模型

MLOps

MLOps或ML Ops是一組旨在可靠有效地在生產中部署和維護機器學習模型的實踐。這個詞是“機器學習”和軟件領域中DevOps的持續開發實踐的組合。

AIOps

指維運越來越依靠演算法（AIgorithmic IT operations的縮寫），不只是AI、機器學習，甚至連傳統大數據分析等，可以輔助IT日常維運工作的演算法都屬之。

Ref: Wikipedia

Ref: AIOps, https://www.ithome.com.tw/news/135082

我想寫個慘字

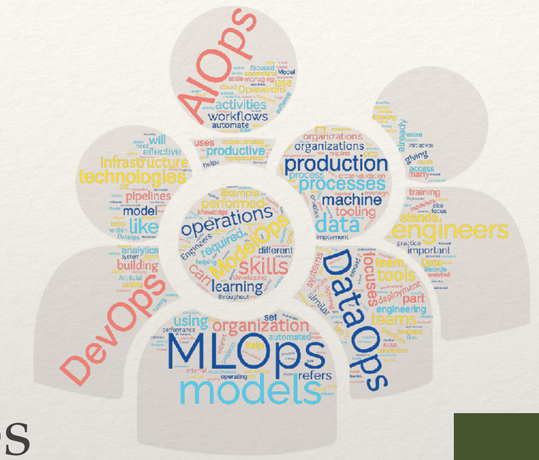
“DevOps 一個殘忍的事實：大部份的 Dev 根本不 care Ops.”

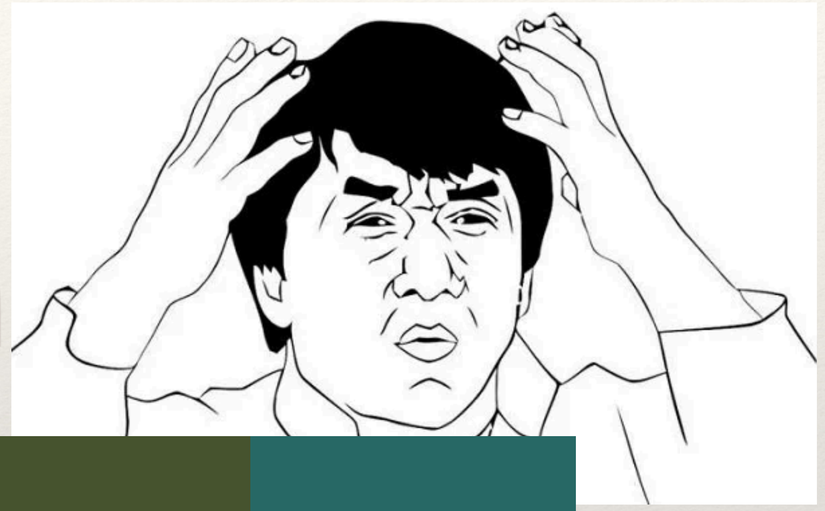
–Dr. Test my code in Production

DevOps 與他的快樂夥伴

AI RD 在哪?

- DevOps is Dev + Ops
- MLOps is **ML model** + Ops
- AIOps is AI for **Ops**
- ModelOps is MLOps + AI + **Governance**

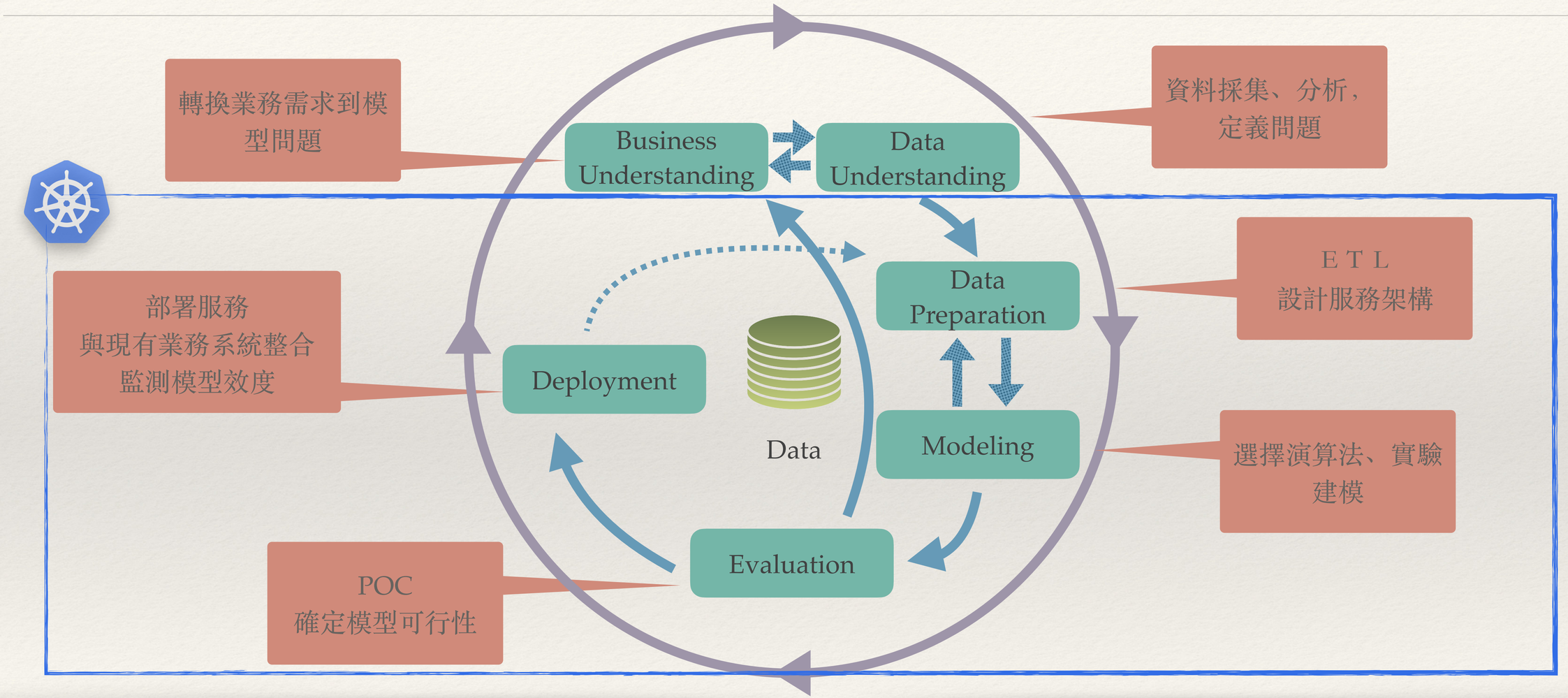




DevOps	ModelOps
MLOps	AIOps

亂啊 ~

AI為核心的解決方案設計準則



Ref: CRISP-DM

AI RD 視角　選擇什麼微服務?

OpenShift Container Platform

Pros

- 原廠 (RedHat) 的支援，24/7 小時服務不間斷
- 穩定、可專注在模型開發
- 有可用的 MLOps 環境
- 模型開發上線維運一條龍

Cons

- 需注意應用程式被平台綁架
- Security Polices 較嚴格
- MLFlow/AutoML/AI 相關套件的安裝限制

AutoKeras
TensorFlow
PyTorch
mlflow

Kubernetes

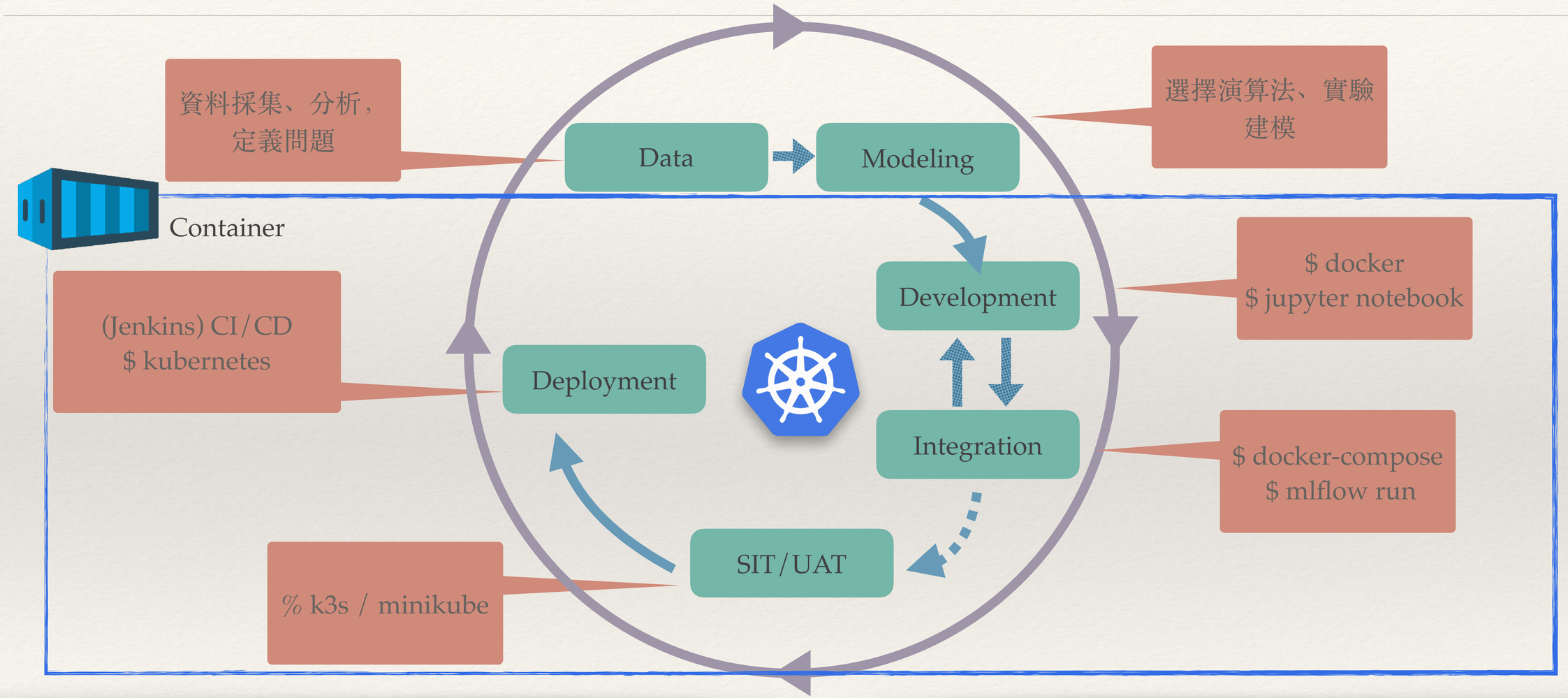
Pros

- 社群的支援，24/7 都可以上社群求救
- 各種 ML/AI 相關軟體或套件安裝都有指引
- 部署原生應用程式容易
- 開發彈性高，porting github 程式不受限

Cons

- 什麼都要自己來
- 安全性容易被忽略，如 0day exploits
- 更新版本需小心相關程式的影響
- 各類模型監控工具需自己想辦法

以部署到 k8s 為目標的開發流程

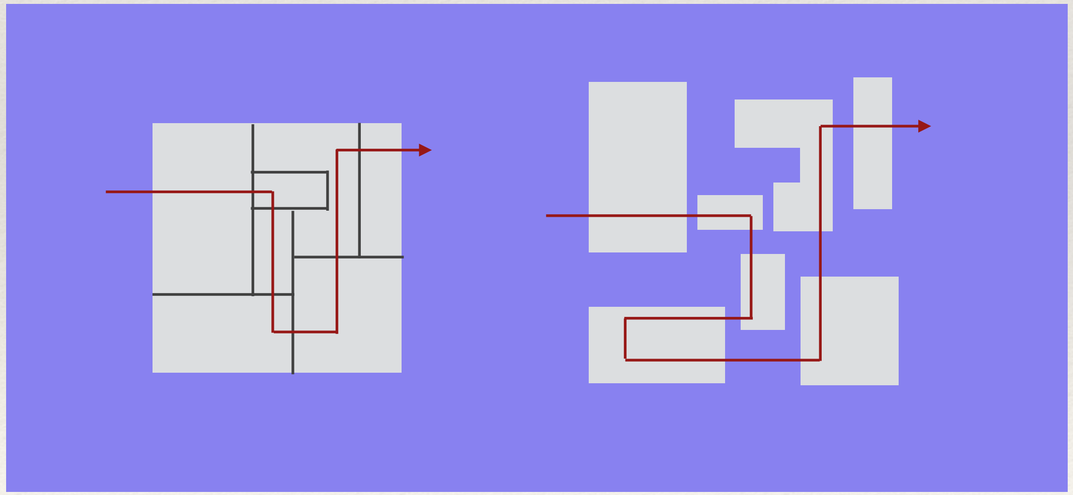


Icons made by Pause08 from www.flaticon.com

(微)服務突然出現異常時....


Honest Status Page
@honest_update
We replaced our monolith with micro services so that every outage could be more like a murder mystery.
7:10 AM · Oct 8, 2015 · Buffer


Trello APP 10:30 PM
Added a new item 加log以便與IT釐清責任 to the checklist 待辦清單 on [Servic 服務建置.



採用微服務，最大的挑戰是
權責歸屬

謎之聲：權責，誰有
權力說是誰的責任

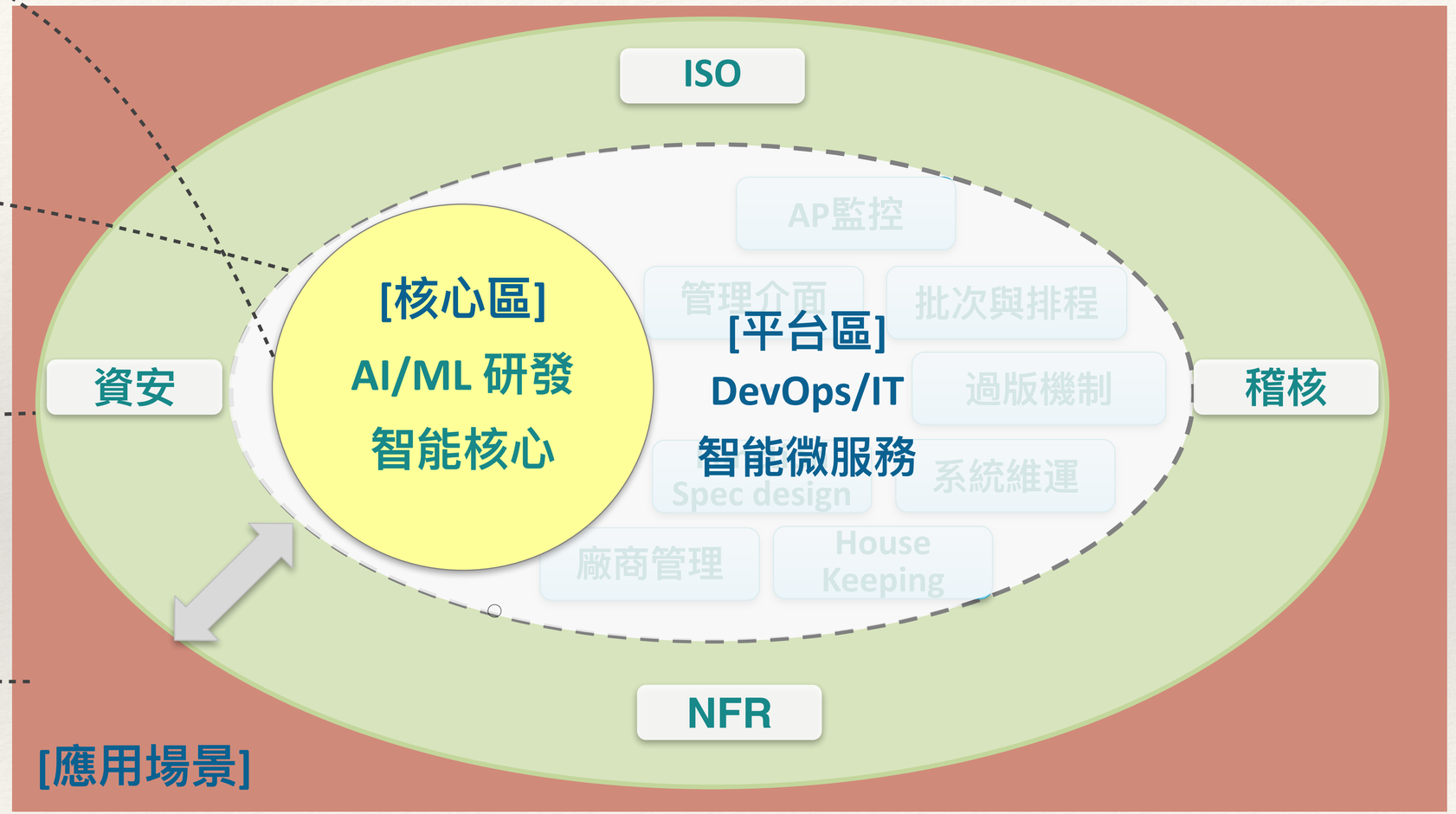
AI服務 - 落地營運 合作角色說明

智能核心
- 演算法研發
- 模型訓練
- 模型迭代優化
- 核心(容器)交付

智能微服務
- AI核心服務建置
- 核心過版
- NFR機制建置
- 系統維運

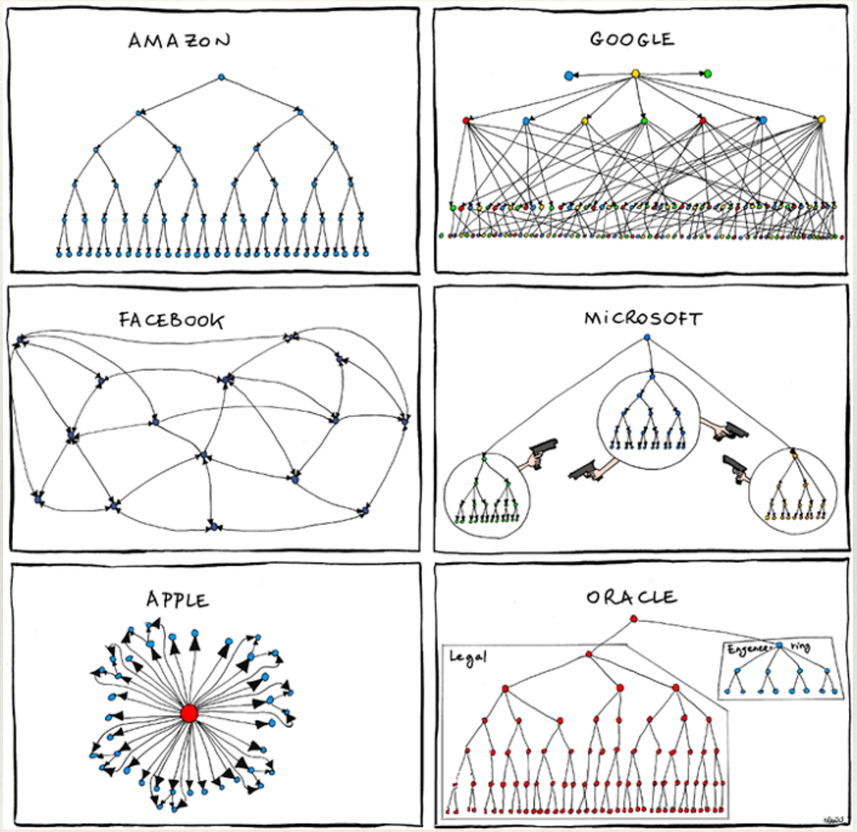
營運規範

各場景系統 AP
- 串接AI服務
- 系統過版
- 系統維運



(跨)微服務的 debug 挑戰

- User 視角：
  - 就是不能用，我怎麼知道是怎樣不能用?
- BU IT 視角：
  - 我這裡的服務沒問題，看 log 是到你們那裡出錯的
- AI IT 視角：
  - 我們這幾次 RC 都沒更版
- AI RD 視角：
  - 我東西之前就給出去了啊，怎麼現在才說有問題?

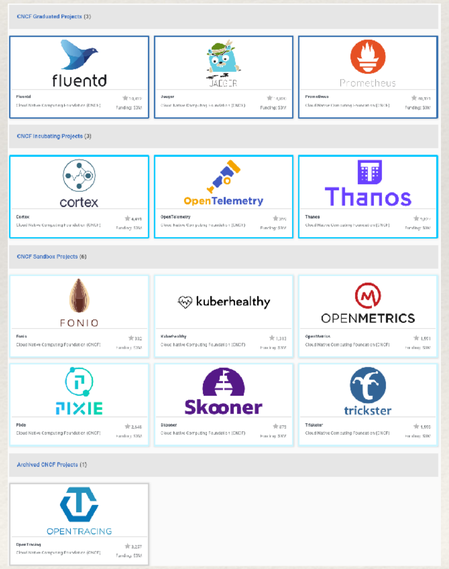


謎之聲：那是 feature

可觀察性（Observability）

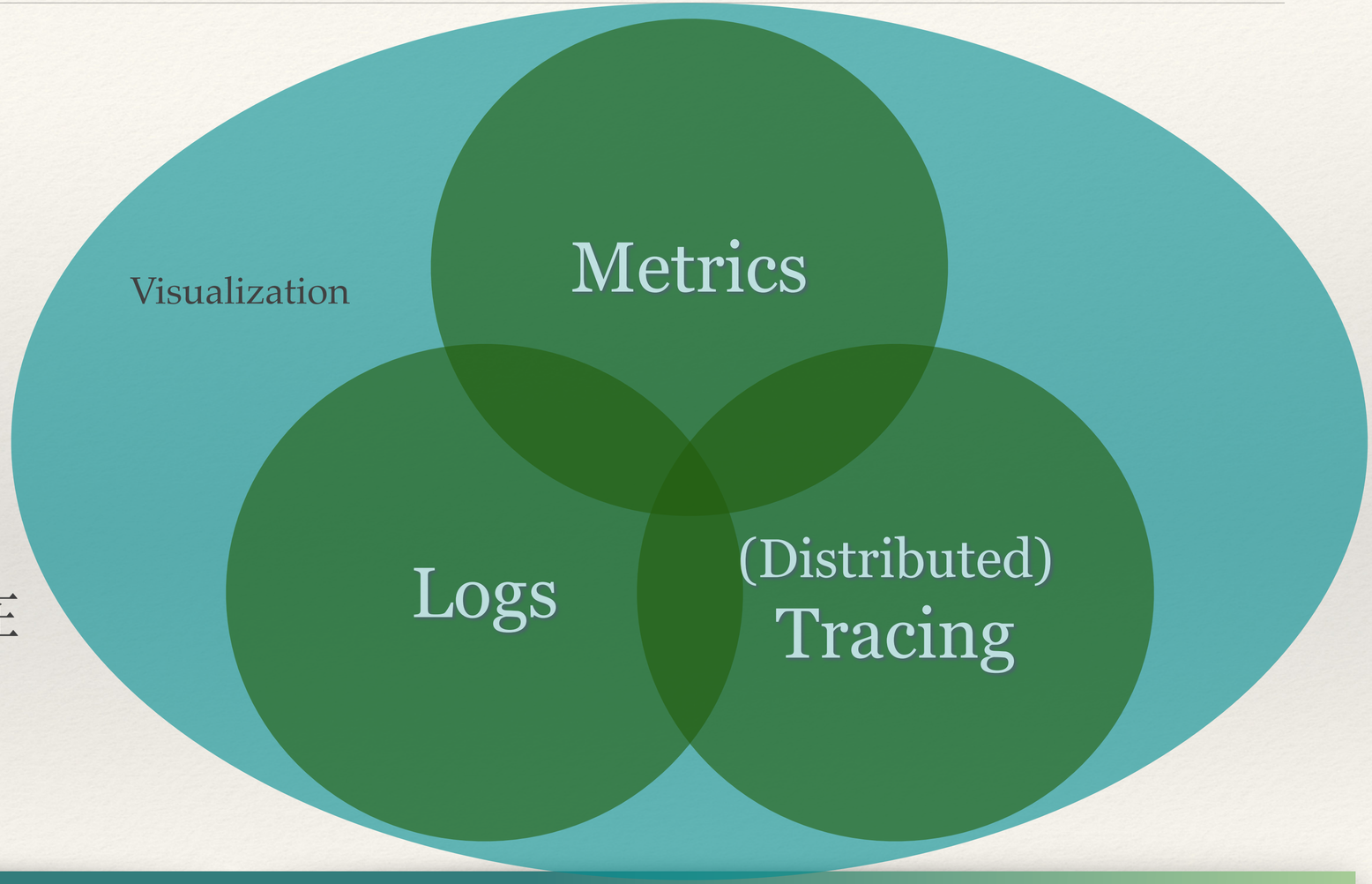
"An observable system is one that exposes enough data about itself so that generating information and easily accessing this information becomes simple."

–Johnny Appleseed



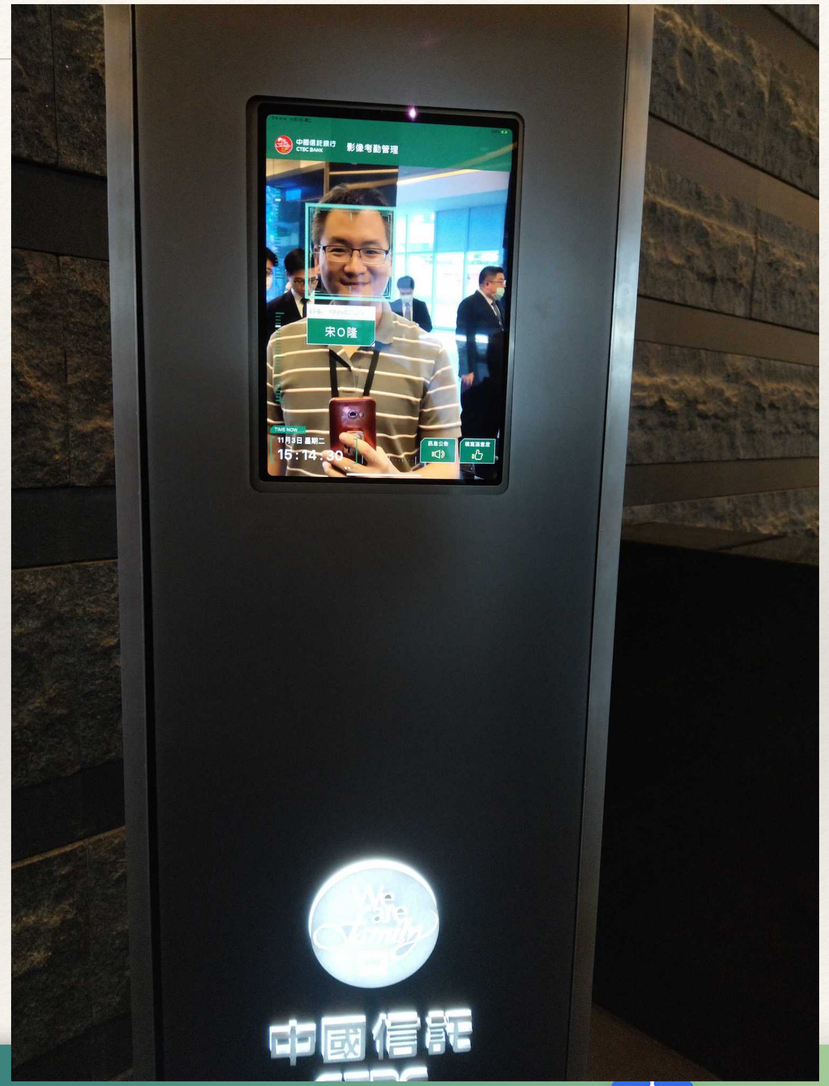
Pillars of Observability

- 程式開發沒有埋好
  traces/metrics/logs
  註定會出現在 #靠北工程師
- 程式開發做好
  traces/metrics/logs
  卻各自為政，也註定會出現在
  #靠北工程師

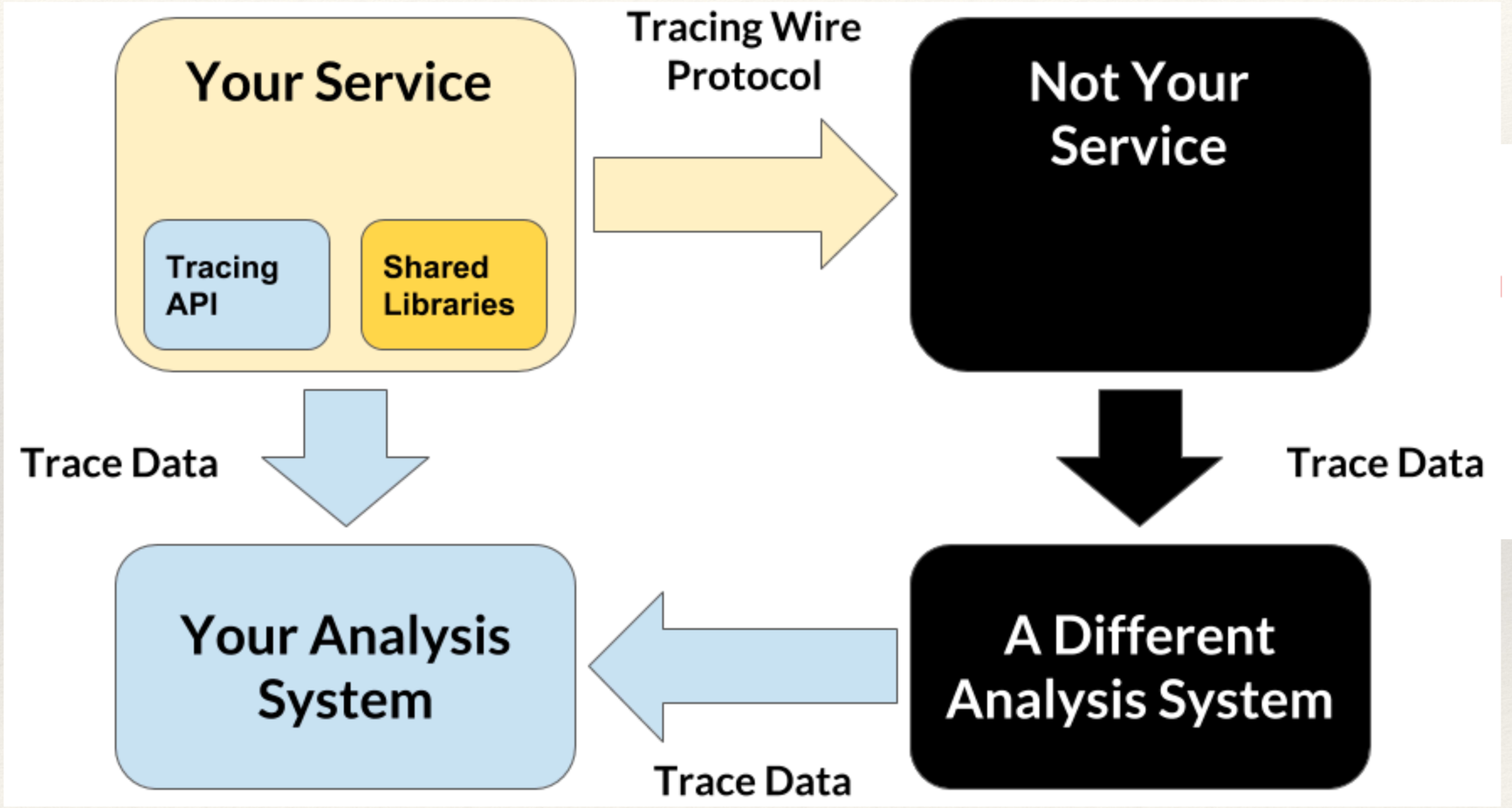


案例

- 問題
  - 有幾個人刷臉打卡特別慢
- 限制
  - 你不是 IT，IT 說大部份都好的 (50%, 95%, 99%?)
  - 核心是你做的，~~你要負責~~
  - 拆機器測，環境就不同了(refer to 在我這裡跑是好的)
  - 看 log ，要先校時 (*why?*)



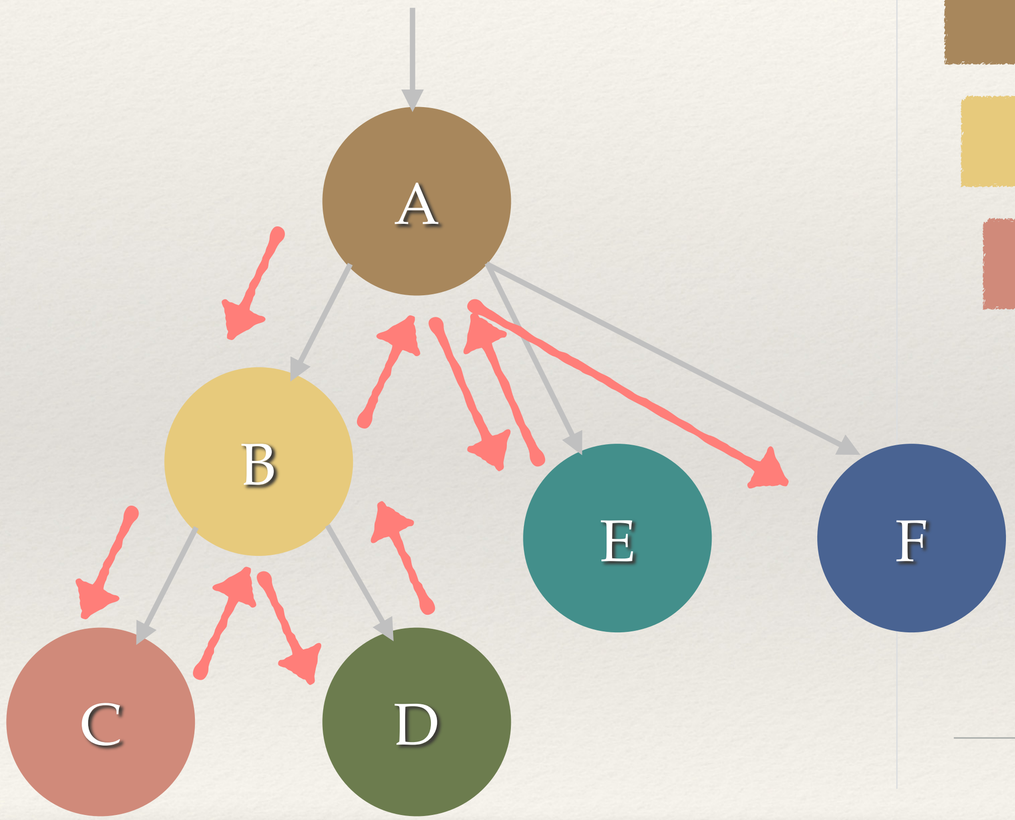
Distributed Tracing

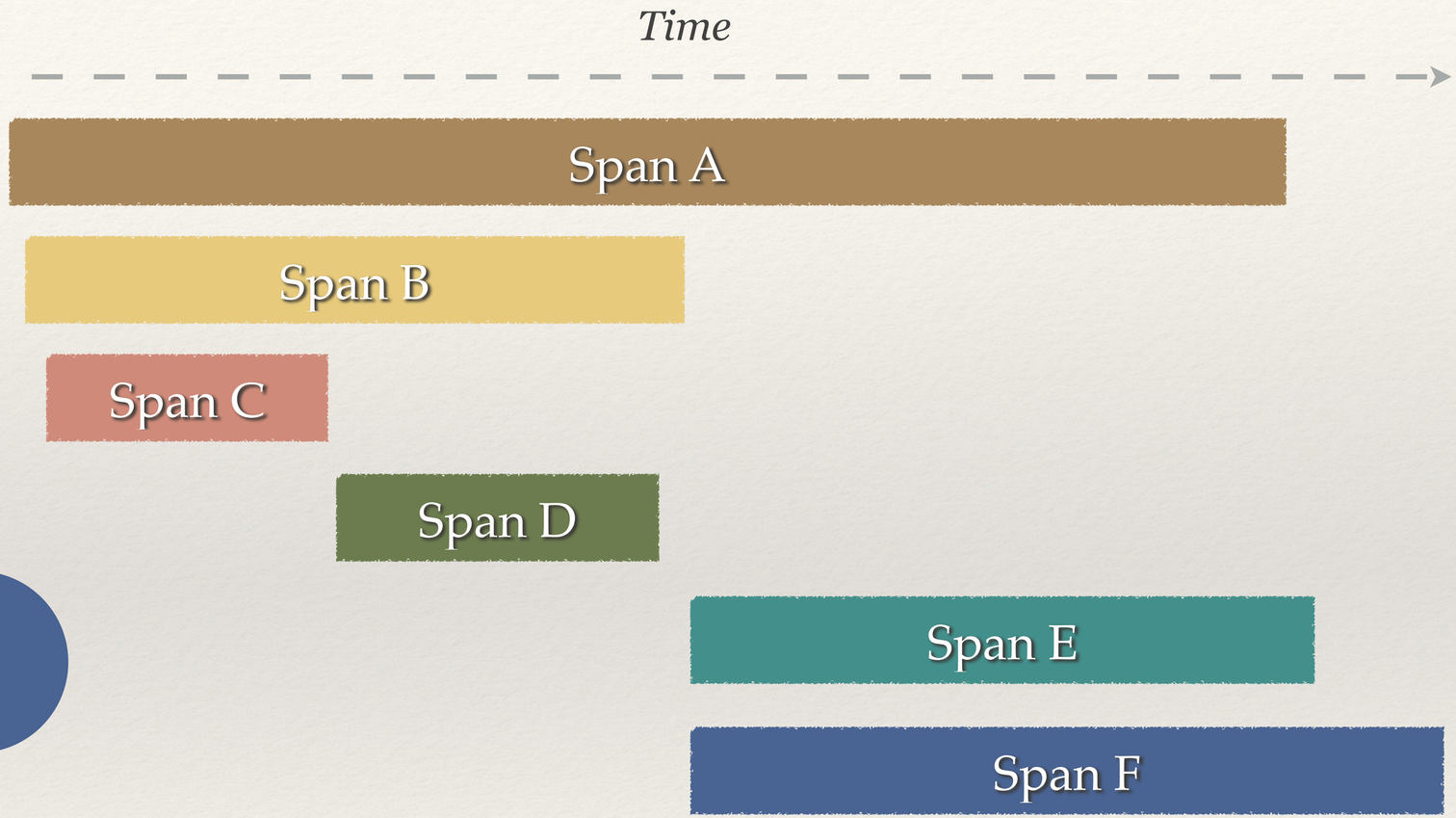


https://opentracing.io/docs/overview/

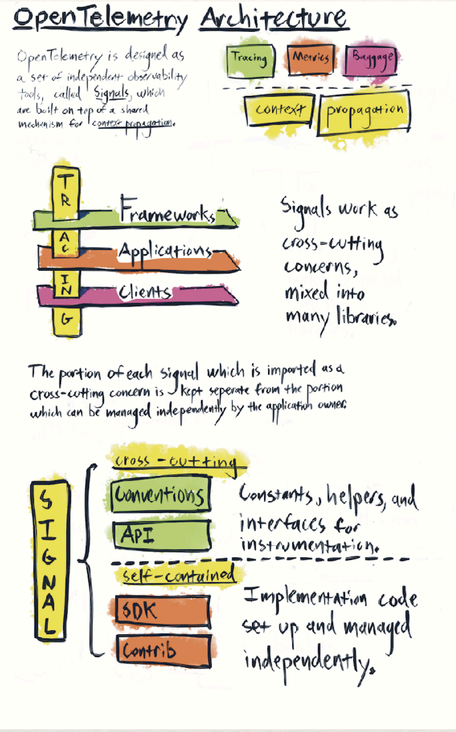
Trace

```
ctx, span := trace.StartSpan(ctx, “start")
defer span.End()
```



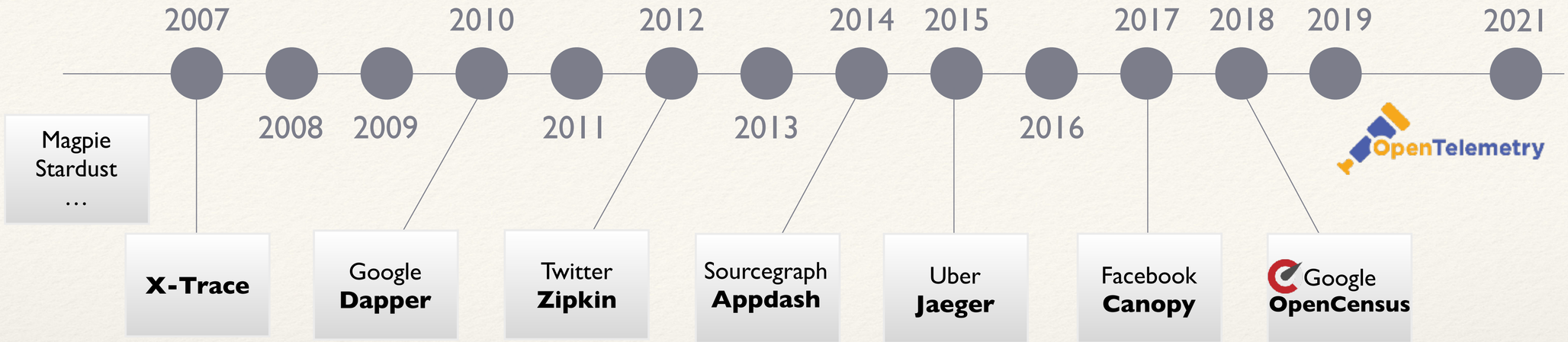


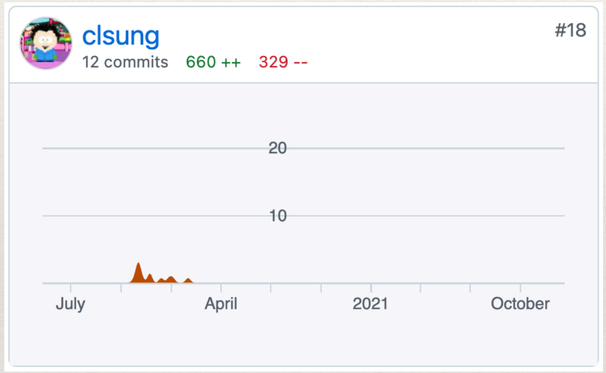
Distributed Tracing 工具選擇



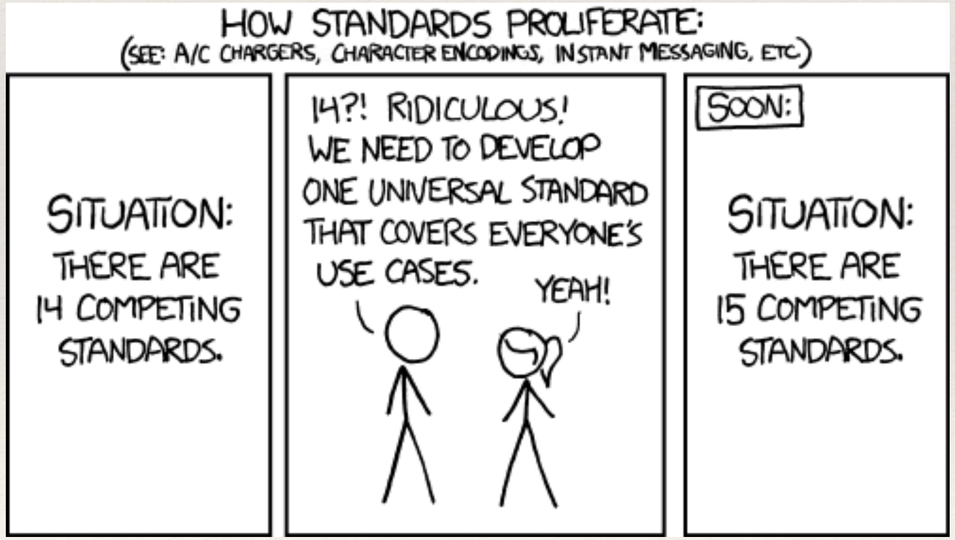
Retired Jaeger Bindings for Python OpenTracing API

OpenTelemetry Python Contrib





Distributed Tracing Frameworks/Tools Timeline



Sampling strategy (OpenTelemetry)

Probabilistic Sampling Processor

```
processors:
  probabilistic_sampler:
    hash_seed: 22
    sampling_percentage: 15.3
```

Tail Sampling Processor

```
processors:
  tail_sampling:
    decision_wait: 10s
    num_traces: 100
    expected_new_traces_per_sec: 10
    policies:
```

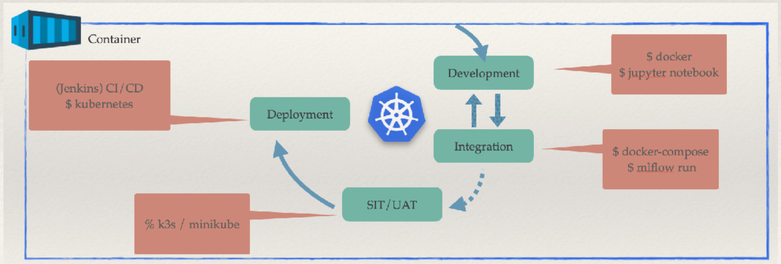
Deprecate the tail-based sampling processor

Deploy strategy (Jaeger)

Sidecar

RD(無感)自己管理

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: oct-deploy
  annotations:
    "sidecar.jaegertracing.io/inject": "true"
spec:
```



DeamonSet

```yaml
apiVersion: jaegertracing.io/v1
kind: Jaeger
metadata:
  name: uat-jaeger
spec:
  agent:
    strategy: DaemonSet
```

多人開發較不佔資源

```yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: ocs-deploy
spec:
      containers:
      - name: xxxx
        image: cbtch/oct:stage-1
        env:
        - name: JAEGER_AGENT_HOST
```

困境

- RD **有一定的比例**易養難教
- RD **有一定的比例**看到新東西，就想換！
- 採用 Open Source
  - **有一定的比例**版本更新後功能不相容 Log4j 表示欣慰
  - **有一定的比例**變化太大，程式碼要重改
- 同上，RD **有一定的比例**已經覺得 Ops 關我啥事了

Open Source
多的是
你不知道的事

總結

- 在中國信託發展 AI 很有趣　　We're hiring: 目前還缺兩名：千軍、萬馬
  - 有很多個場景，在甲方裡當乙方
  - 所以才需要微服務
- 以 AI/ML 工程師或是資料科學家的角度來看 Kubernetes 微服務幫助了我們什麼
  - 開發核心 —> 開發可上線的核心
  - ML 相關套件愈來愈多，建立開發環境也變得容易
- 可觀察性
  - 讓我們避免瞎子摸象，碰不到的系統也能抓蟲
  - 另一方面，還在發展中的 open source ，也造成了學習曲線

Thank you!