Big Health Data and Precision Medicine/Health
健康大數據與精準健康醫療
運用電子病歷整合醫療資訊之挑戰

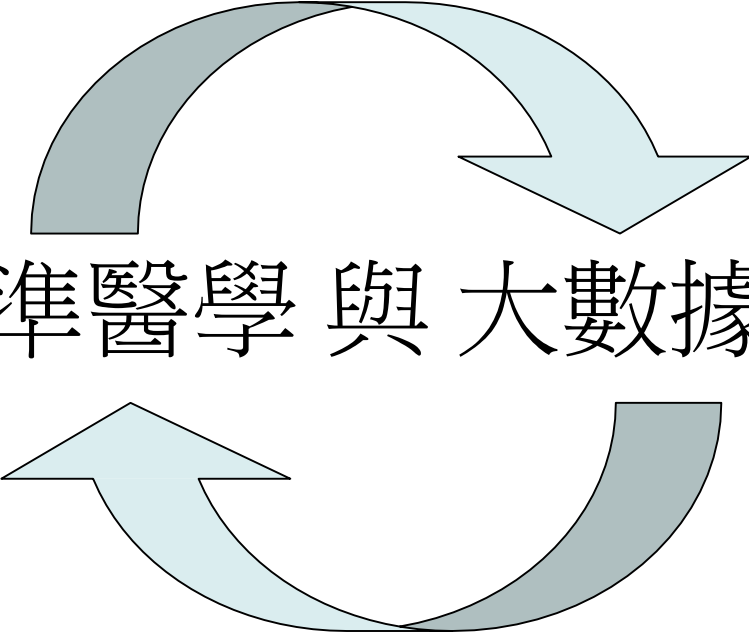Chien-Yeh Hsu 徐建業 PhD

臺北護理健康大學資訊管理系所

National Taipei University of Nursing and Health Sciences

臺灣醫學資訊學會

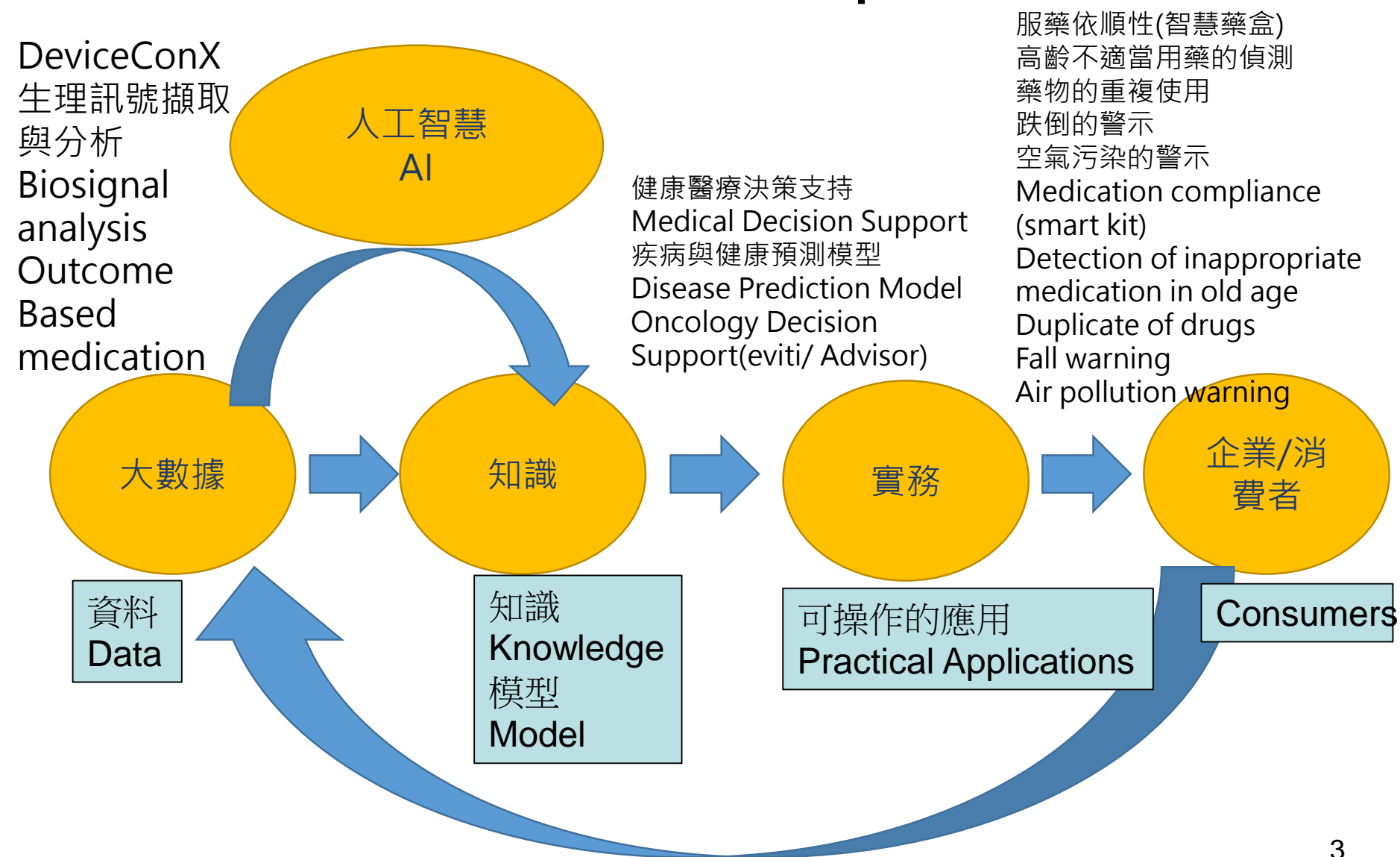Taiwan Association for Medical Informatics TAMI

臺灣電子健康學會

Taiwan Electronic Health Association - TEHA

精準醫學 與 大數據

智慧健康照護與醫療大數據之發展流程
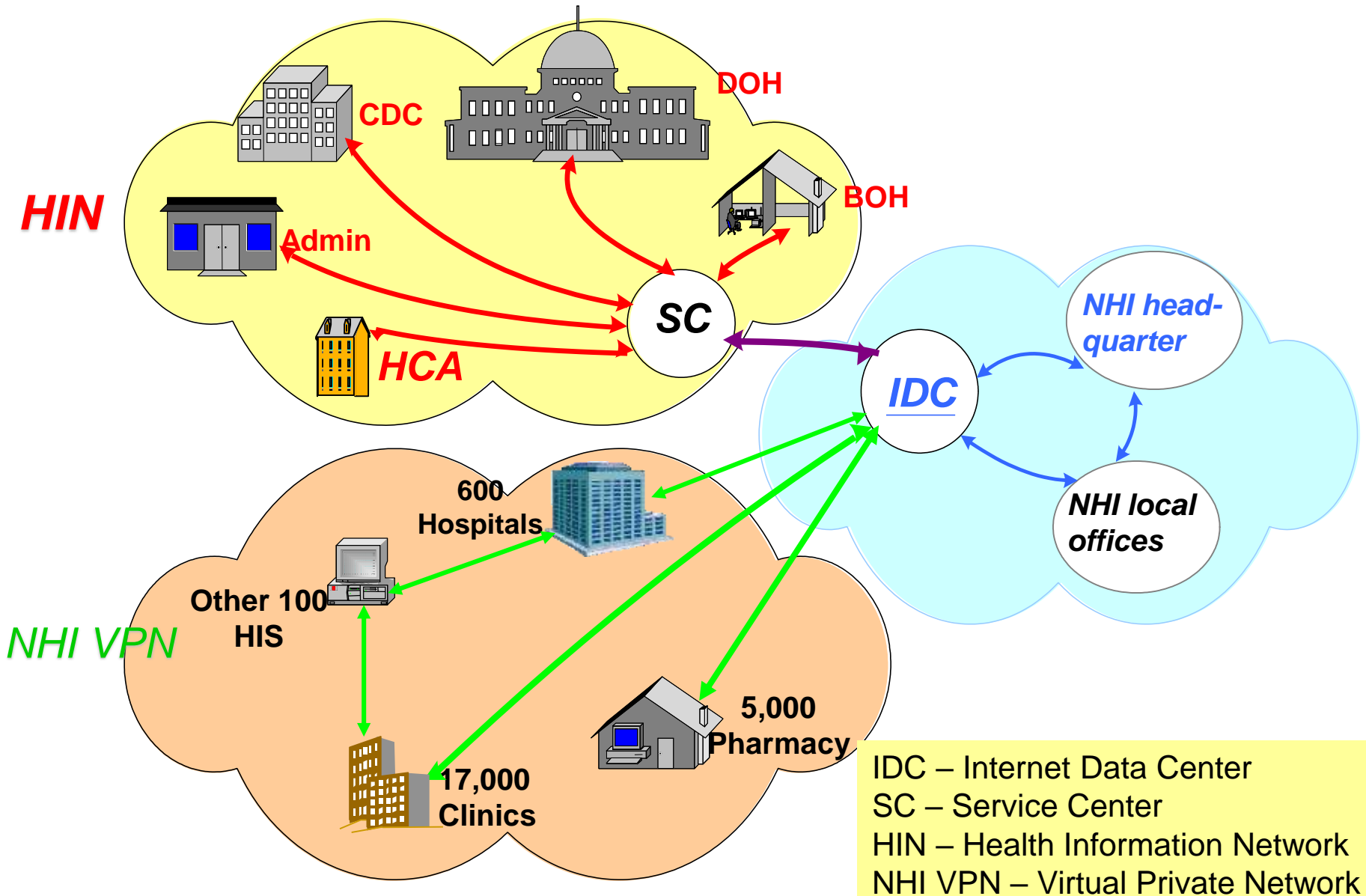## Smart Health care with the use of big health data for Precision health care and personal care

DeviceConX
生理訊號擷取
與分析
Biosignal
analysis
Outcome
Based
medication

人工智慧
AI

健康醫療決策支持
Medical Decision Support
疾病與健康預測模型
Disease Prediction Model
Oncology Decision
Support(eviti/ Advisor)

服藥依順性(智慧藥盒)
高齡不適當用藥的偵測
藥物的重複使用
跌倒的警示
空氣污染的警示
Medication compliance
(smart kit)
Detection of inappropriate
medication in old age
Duplicate of drugs
Fall warning
Air pollution warning

大數據

知識

實務

企業/消費者

資料
Data

知識
Knowledge
模型
Model

可操作的應用
Practical Applications

Consumers

3

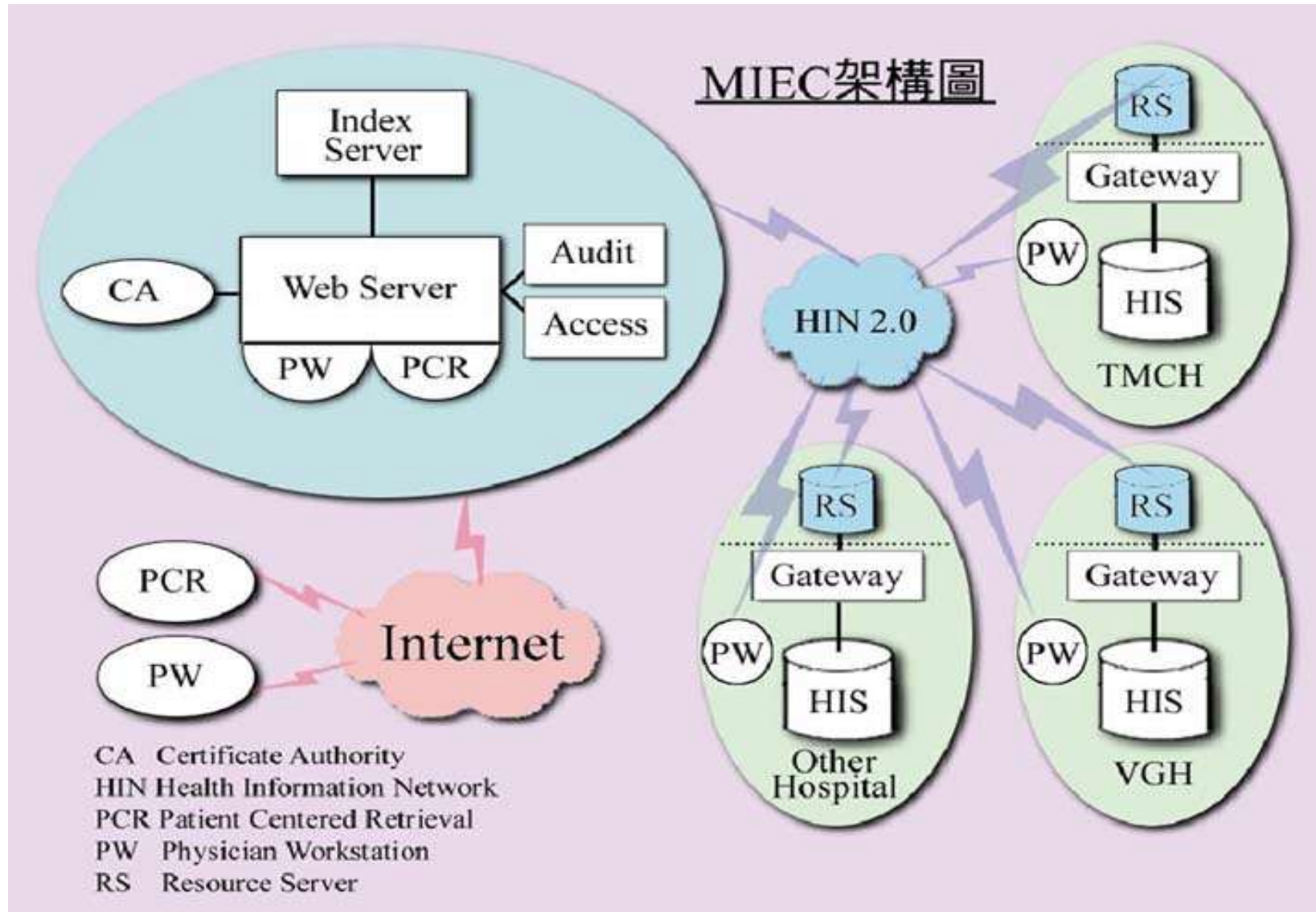# Major Trends of Medical Informatics
〈Everything is based on Data〉

- Virtualization (digitization)
    - of medical data for education/service/research

- Aggregation (integration)
    - of biomedical information and systems

- Standardization & Exchange (of EMR)
    - of medical information and healthcare management/practice

- Intelligence (smart)
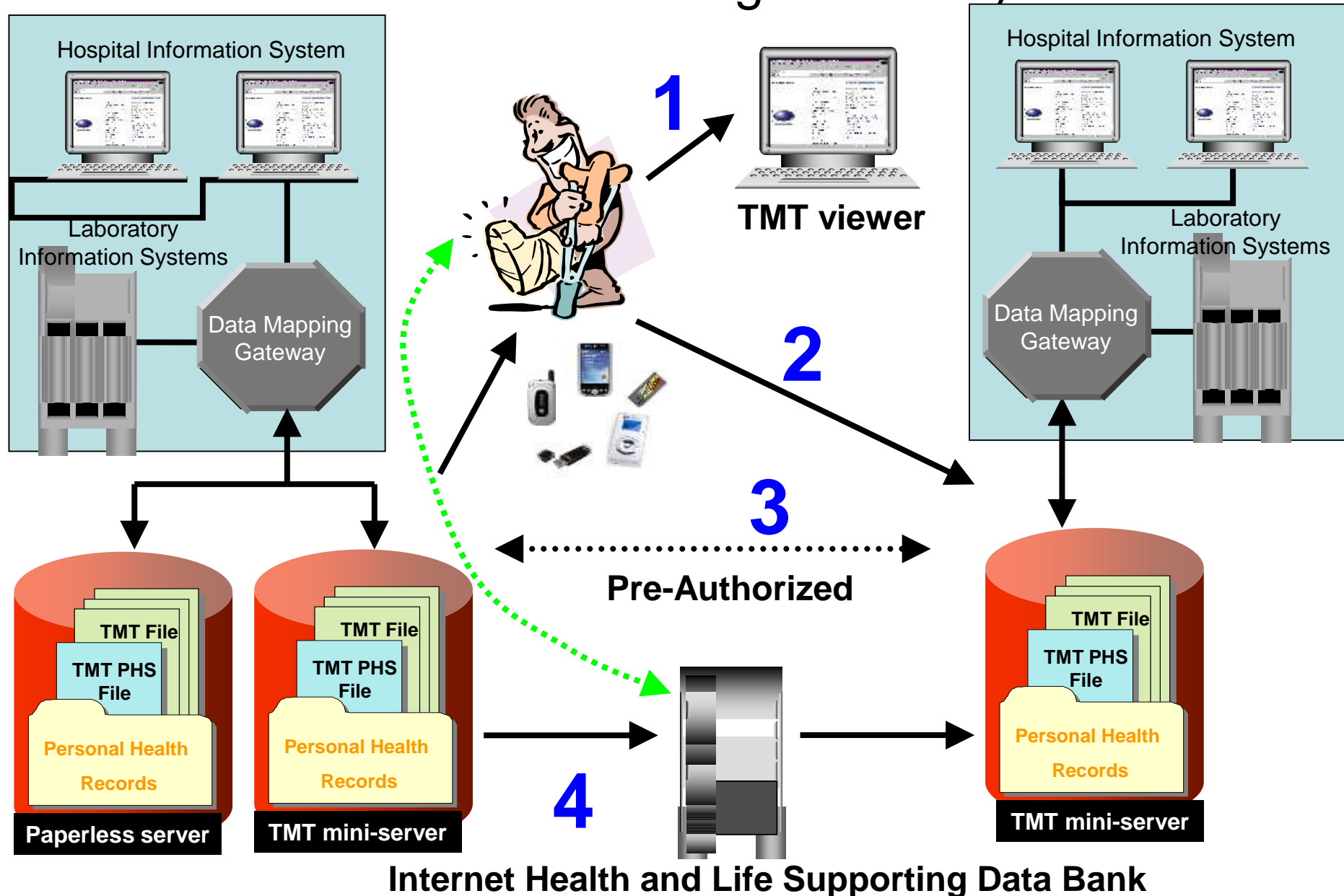   - for knowledge-based healthcare

The NHI 臺灣建保資料 VPN

HIN

CDC

DOH

BOH

Admin

SC

HCA

NHI VPN

600 Hospitals

Other 100 HIS

17,000 Clinics

5,000 Pharmacy

IDC

NHI head-quarter

NHI local offices

IDC – Internet Data Center
SC – Service Center
HIN – Health Information Network
NHI VPN – Virtual Private Network

Medical Information Exchange Center – MIEC 2000

個人化健康資訊整合架構
TMT File Exchange Pathway

Internet Health and Life Supporting Data Bank

行政院衛生署

ＥＥＣ環境建置架構圖 EMR Exchange Center

EMR providing Hospitals

National EEC Center

EMR Reading Hospitals

EEC 閘道器 (Gateway)

提供病歷醫院

全國電子病歷交換中心主機 (EEC)

EEC 閘道器 (Gateway)

調閱病歷醫院

DICOM

EMR

PACS

跨院調閱同意書

1. 查詢
2. 調閱
3. 下載

Download, Querying, and Reading
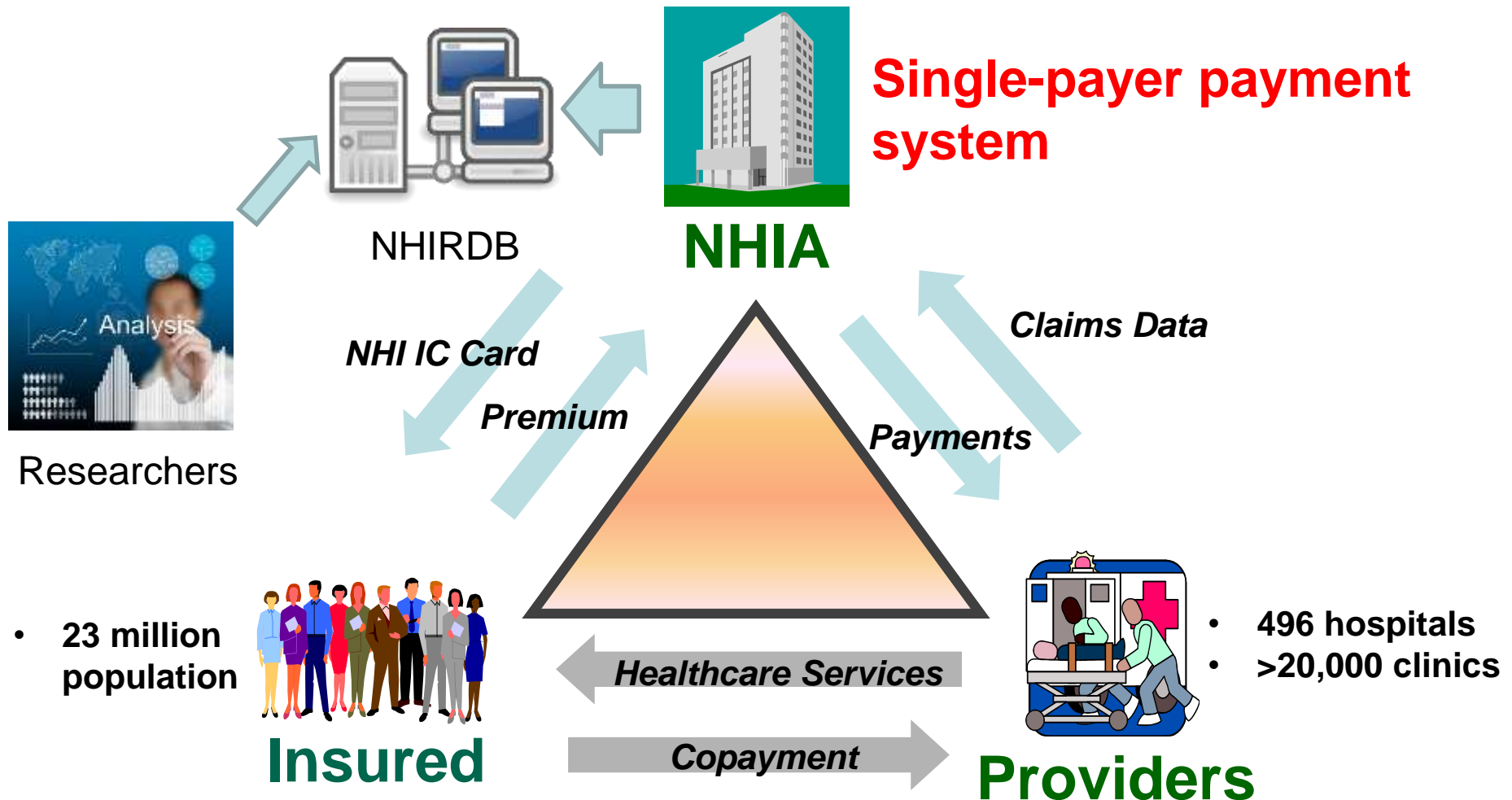
Ministry of Health
Image Exchange Center 衛生署

影像交換平台

Health Insurance Center
健保局

※健保IC卡執行
CT,MRI,PET就醫資料
（24小時上傳）

傳送索引

批次傳送索引

索引資料庫

Index Server

**134 hospitals,
2010-2011
upload index 2,168,063
request: 6,592
download: 81,108**

甲醫院 Hospital A

Hospital B

乙醫院

HIS
報告

PACS
影像

傳送報告與影像

Gateway

健保VPN

健保VPN

Gateway

傳送報告與影像

PACS
影像

HIS
報告

eSignature
加簽機構章

Radiology
放射科

影像報告
資料庫

Image
Report
Database

Image
Report
Database

影像報告
資料庫

加簽機構章

Radiology
放射科

下載影像

下載影像

Download image

醫令系統 CPOE

CPOE 醫令系統

Download image

查詢索引/調閱影像

醫師卡+健保卡
與簽署同意書

醫師卡+健保卡
與簽署同意書

查詢索引/調閱影像

Request for image

Dual Card system and
inform consent

Request for image

Dual Card system and
inform consent

10

Taiwan National Health Insurance

Single-payer payment system

NHIRDB

NHIA

Researchers

NHI IC Card

Premium

Claims Data

Payments

- **23 million population**

Healthcare Services

Copayment

- **496 hospitals**
- **>20,000 clinics**

Insured

Providers

臺灣研究生物醫學的優勢: 全世界唯一的全民建保資料庫

http://w3.nhri.org.tw/nhird/

32

NHIRDB in Taiwan

- NHIRDB (National Health Insurance Research Database)
  - 12 years of de-identified claim database for 23 million people
  - Cohort DB (Five 1-million people groups for 13 years)
  - Disease-specific DB (16 disease groups)
  - Random sample DB (outpatient 1/500, inpatient 1/20)
  - generates >100 research papers a year

≫Data files

Each year, BNHI collects data from the National Health Insurance program and sorts it into data files, including registration files and original claim data for reimbursement. These data files are de-identified by scrambling the identification codes of both patients and medical facilities and sent to the National Health Research Institutes to form the original files of NHIRD.

The registration files include:
- Registry for contracted beds (BED)
- Registry for contracted specialty services (DETA)
- Registry for contracted medical facilities (HOSB)
- Supplementary registry for contracted medical facilities (HOSX)
- Registry for board-certified specialists (DOC)
- Registry for medical personnel (PER)
- Registry for catastrophic illness patients (HV)
- Registry for medical services (HOX)/td>
- Registry for drug prescriptions (DRUG)
- Registry for beneficiaries (ID)

The Original Claim Data include:
- Monthly claim summary for inpatient claims (DT)
- Monthly claim summary for ambulatory care claims (CT)
- Inpatient expenditures by admissions (DD)
- Details of inpatient orders (DO)
- Ambulatory care expenditures by visits (CD)
- Details of ambulatory care orders (OO)
- Expenditures for prescriptions dispensed at contracted pharmacies (GD)
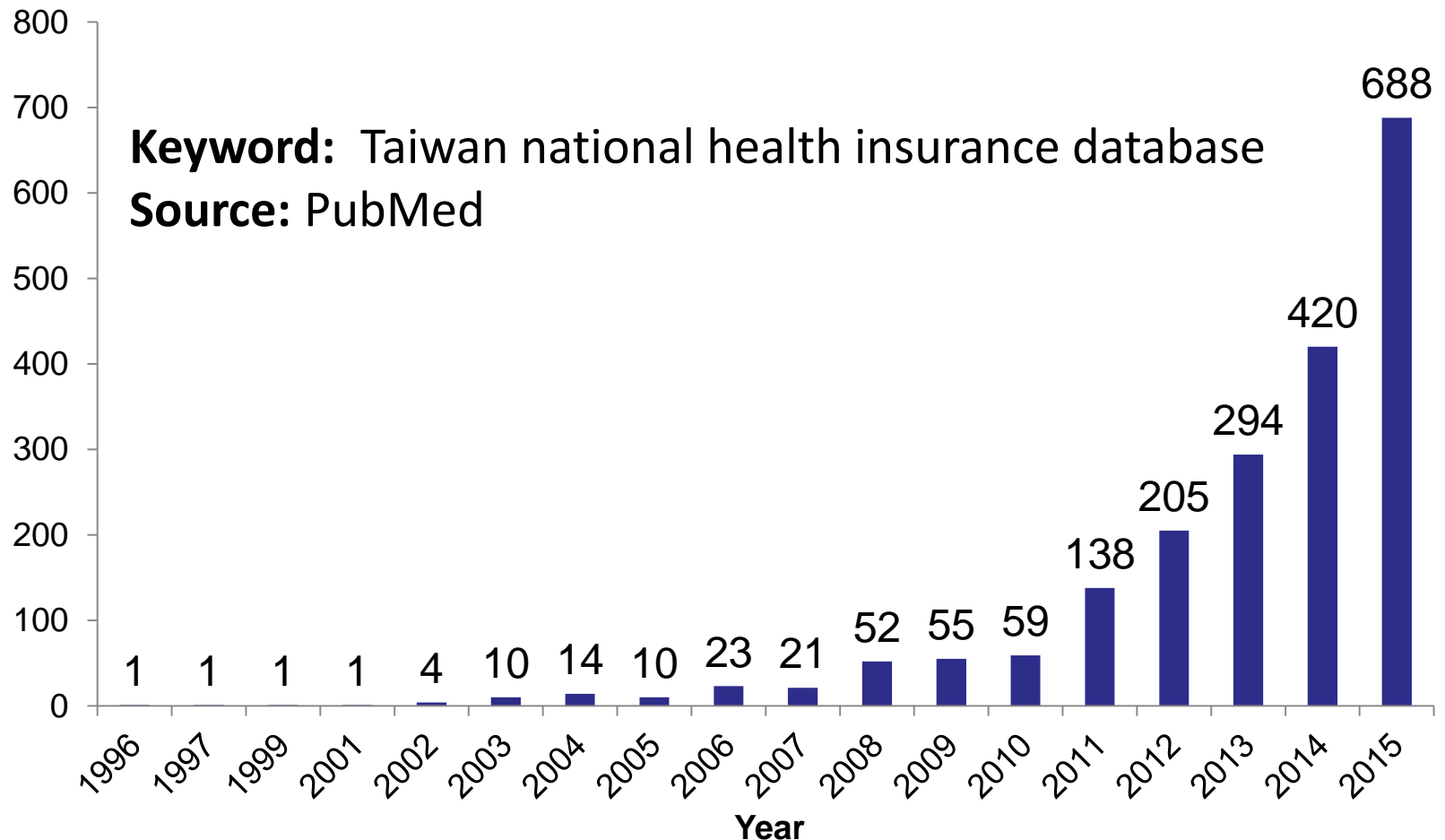- Details of prescriptions dispensed at contracted pharmacies (GO)

NHRI 全民健康保險研究資料庫
National Health Insurance Research Database
ENGLISH

Papers Published by Using NHI Research Database (1996-2015)

Keyword: Taiwan national health insurance database
Source: PubMed

At least 1 paper published per day since 2014

臺灣人體生物資料庫透過 *Taiwan view*網站，已開始釋出基因型 相關資訊

Taiwan BioBank 宗旨

臺灣人體生物資料庫成立目的是結合生活習慣、環境因子與生物標誌等資訊，建立屬於臺灣本土的人體生物資料庫，為生物醫學研究蒐集龐大的生物檢體與健康資訊，提供國內各領域的研究學者申請使用。

自2003年全世界第一例全基體定序結果公佈後，全世界的生物醫學研究藉由這結果有了大規模的進步。 而臺灣人體生物資料庫希望透過建立全基因體相關資料，協助生物醫學研究學者找出對抗疾病的發生，進展和治療上的線索，促進國人未來的健康。

本資料庫採用全基因體定型與全基因體掃描兩個策略建立基因體資料，資料之釋出將去除身份識別資訊， 以確保達到保護參與者隱私權益，提供各領域的研究學者申請使用，進行描述性分析，作為基因醫學研究的基礎，以建立研究之初步假說。

【全基因體定序】

本資料庫同時選用Ion Proton與Illumina平台，透過血液中萃取的DNA檢體進行次世代全基因體的定序， 將會有助於全面性的了解臺灣國民基因遺傳變異以及在基因層次上健康與疾病的關係，將為國內生物醫學研究上帶來幾點助益：

1. 建立臺灣自己的參考序列
2. 為臺灣建立健康對照組的序列資訊
3. 作為基因填補法(genetic imputation)的模板，增加研究效益
4. 提供台灣健康族群低頻率變異(rare allele)分布情形
5. 有助於發展全基因體關聯性研究(genome-wide association study)

About   Browse Data   Search   Download   Tutorial   Data application   Registration   Login

https://taiwanview.twbiobank.org.tw/index

3

電子病歷之定義及背景資料
EMR Definition and background

Electronic Medical Record (EMR)

- A generic term used to describe **computer-based patient medical records**. Also includes systems which keep track of other relevant medical information.

- An **electronic patient record** that resides in a system specifically designed to support users through availability of **complete and accurate data,** practitioner **reminders and alerts, clinical decision support** systems, link to bodies of **medical knowledge,** and other aids.

- **patient's personal details + other functions**

Electronic Health Record (EHR)

- *An Electronic Health Record (EHR) is any information relating to the past, present or future physical and mental health, or condition of an **individual** which resides in electronic system(s) used to capture, transmit, receive, store, retrieve, link, and manipulate multimedia data for the primary purpose of **providing health care and health-related services.***

  - *(Murphy, G. F. et al. (1999), Electronic Health Records: Changing the Vision, Harcourt Brace & Co., Philadelphia, p.5.)*

電子病歷
Levels of an Electronic HealthCare Record (EHCR)

- **Automated Medical Record** (AMR): a paper-based record with some computer-generated documents.

- **Computerized Medical Record** (CMR) : makes the documents of AMR electronically available.

- **Electronic Medical Record** (EMR) : restructures and optimizes the documents of AMR and CMR ensuring inter-operability of all documentation systems.

- **Electronic Patient Record** (EPR) is a patient-centered record with information from multiple institutions.

- **Electronic Health Record** (EHR) adds general health related information to the EPR.

Medical Records Institute

實施與利用電子病歷面臨之挑戰

- 適法性Regarding to the regulations and policies
  - 醫療法(medical law, medical service act)規定病歷需書面並至少保留十年(saved for 7 years)
  - 個人資料保護法(Personal Data Protection Law)第 6 條
  - 有關病歷、醫療、基因、性生活、健康檢查及犯罪前科之個人資料，不得蒐集、處理或利用。但有下列情形之一者，不在此限…
  - 醫療機構電子病歷製作及管理辦法
- 既有之龐大數量實體病歷如何處理Importing the exist records into a new architecture
  - 包括病歷、X光片及其他檢查記錄(different kinds records)
  - 大量影像儲存與傳輸速度 (image)
  - 影像壓縮儲存技術尚有爭議 (image compression)
  - 大量門診調片需求傳輸速度將是關鍵(transmission speed)
  - 與HIS結合應用軟體 (integrated with the current HIS)

實施與利用電子病歷面臨之挑戰

- 病歷醫令輸入及查閱之方便性是否適當User Interface
  - 諸如醫學圖型、影像運用工具熟悉及方便使用(software for medical imaging)
  - 大部份醫院的門住診醫囑系統各自開發，除部分檢查(驗)使用共同資料倉儲外，其餘尚待聯結，無法以病人為中心貫穿(system integration for HIS)
- 資料標準化問題Standards of information
  - 國內醫療資料代碼雖有健保資料交換標準，但僅提供醫令部分，其餘尚未統一(health insurance codes is ok, but others?)
- 資料之安全管理Security issues
  - 電子簽章(Electronic Signatures Law and the technology)技術尚未大眾化，相關法令亦未突破，電子病歷合法性尚有爭議

Medical Information Standards

- Logical Observations, Identifiers, Names and Codes (LOINC)
- Diagnosis-Related groups (DRG)診斷關聯群
- Current Procedural Terminology(CPT)程序
- International Classification of Diseases (ICD)
- North American Nursing Diagnosis Association International(NANDA-I)
  - Nursing Interventions Classification(NIC)
  - Nursing Outcomes Classification(NOC)
  - International Classification of Nursing Practice(ICNP)
- Systematized Nomenclature of Medicine (SNOMED)病理
- Medical subject Headings (MeSH)
- Unified Medical Language System (UMLS)
  - RxNorm

Medical Information Standards

- Digital Imaging and Communications in Medicine(DICOM)影像

- Health Level Seven(HL7)
    - Clinical Document Architecture(CDA)
    - Fast Healthcare Interoperability Resources(FHIR)

- Integrating the Healthcare Enterprise(IHE)
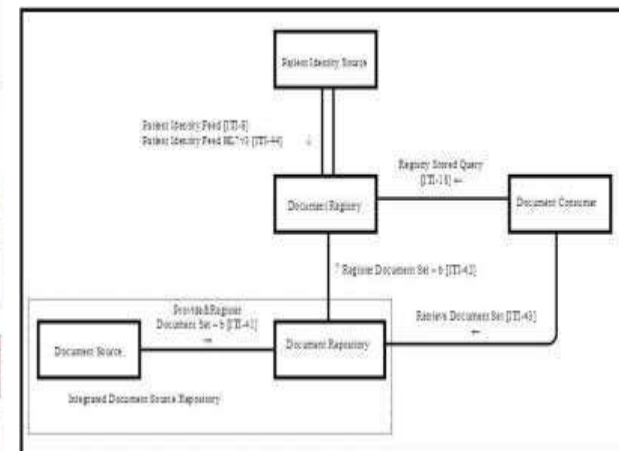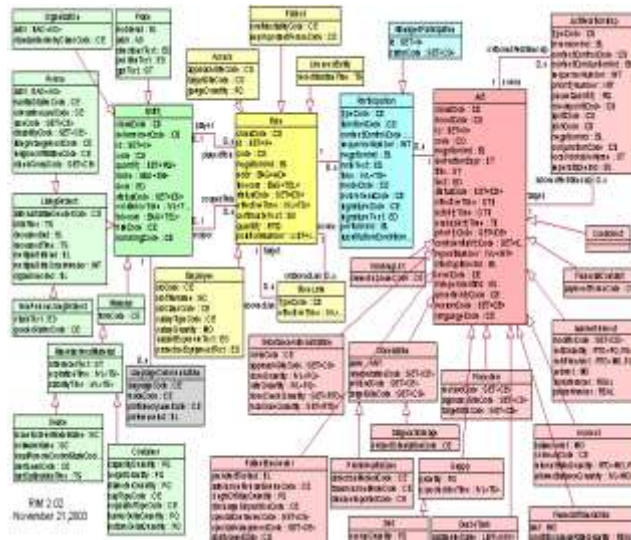
Enforcing interoperability

• Interoperability

"*Ability of two or more systems or components to exchange information and to use the information that has been exchanged*"

Functional interoperability

Semantic interoperability

IEEE Standard Computer Dictionary, IEEE, 1990

跨院整合之
個人化電子病歷與資料中心

- Medical Information Exchange Center and Clinical Data Repository
  - consistent personal medical record across hospitals
- Many issues
  - Security/Privacy/Confidentiality
  - Standardized exchange format
  - Architecture that minimize the impact to current HIS (hospital information system)

Data Architecture by Subjects with Active Analysis and Assessment
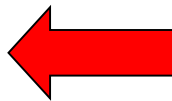
Integrated Data Center for Bio-medical Informatics

Health Insurance DB

Provide Data
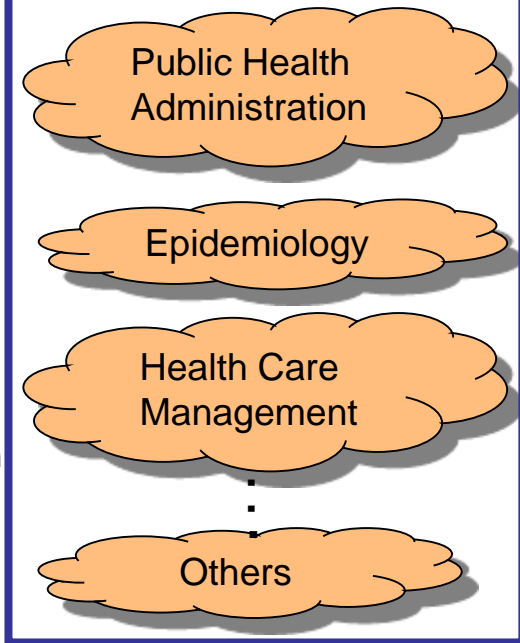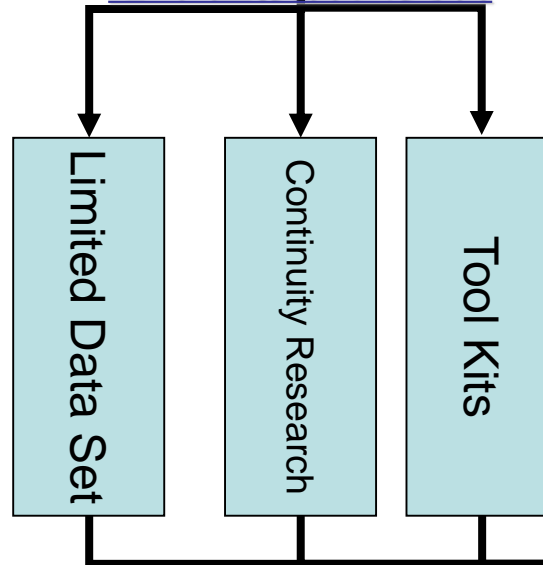
Cause of Death DB

Cancer Register DB

Household (census) Register DB

Redundant / Feedback

Regular Data Released

De-identification

Public Health Administration

Epidemiology

Health Care Management

:

Others

未通過專法，本中心不直接釋出資料

Limited Data Set

Continuity Research

Tool Kits
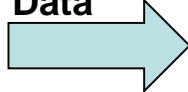
Data Released after IRB approval
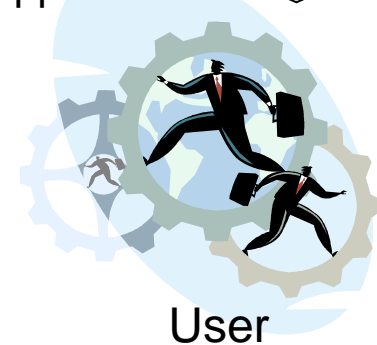
User
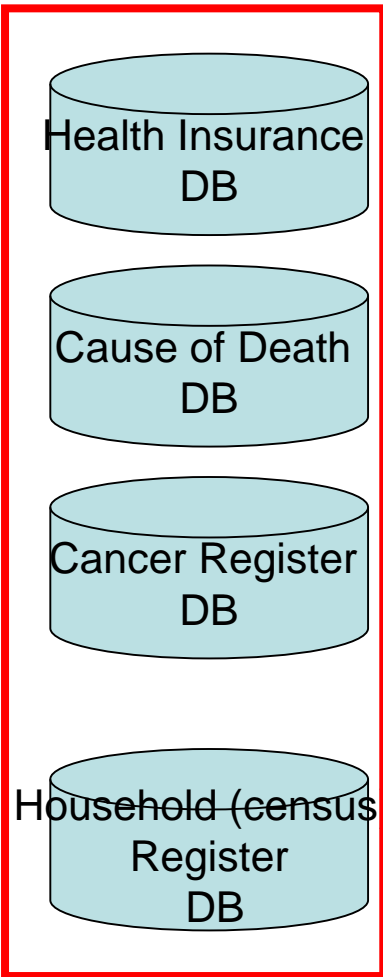
Scenario A

Center of Health Information Application 2009
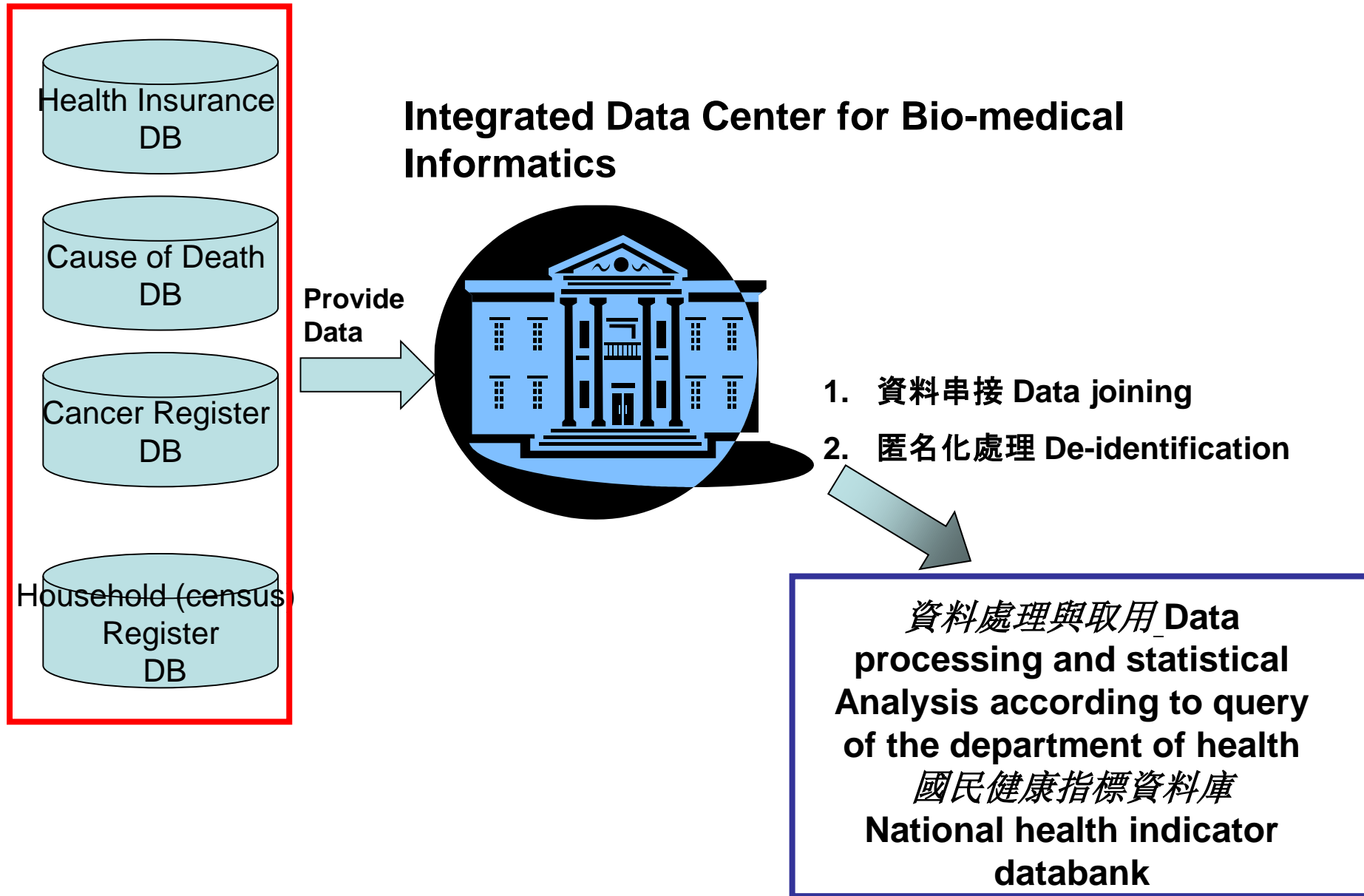
Health Insurance DB

Cause of Death DB

Cancer Register DB

Household (census) Register DB

Provide Data

Integrated Data Center for Bio-medical Informatics

1. 資料串接 Data joining

2. 匿名化處理 De-identification

資料處理與取用 **Data processing and statistical Analysis according to query of the department of health** *國民健康指標資料庫* **National health indicator databank**

Taiwan Information Integrated Platform for Cancer Research : cancer Data Standards Repository and Grid Computing
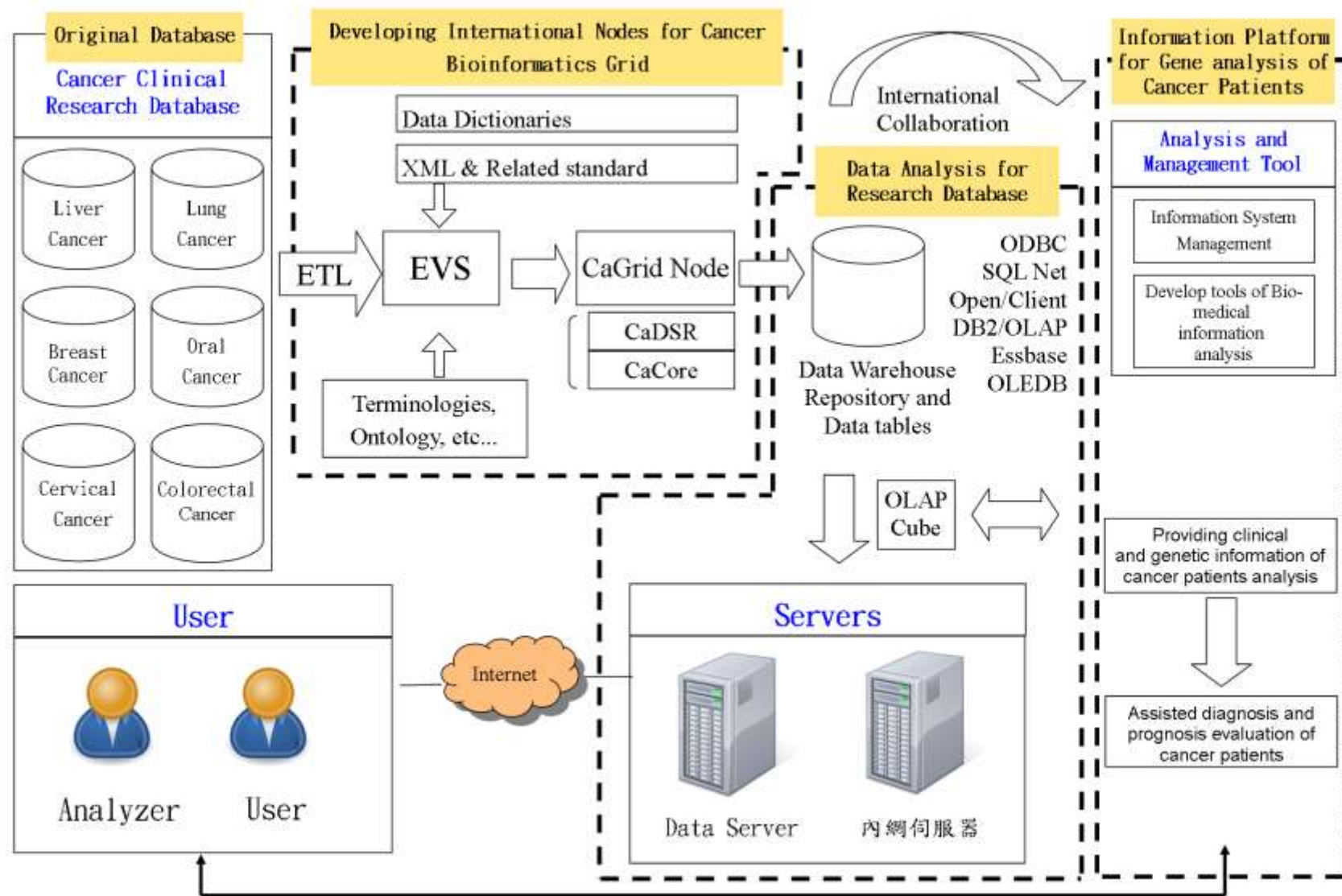
Functions of PHR
個人健康資料的功能

- easy look up of <u>patient data</u> by doctors and nurses
- easy claims processing by <u>insurance</u> companies, archival and legally acceptable records
- easy <u>transfer and share information</u> when a person changes medical facilities
- easy access to the <u>medical/health history</u> of a patient
- providing material for <u>research and studying</u>
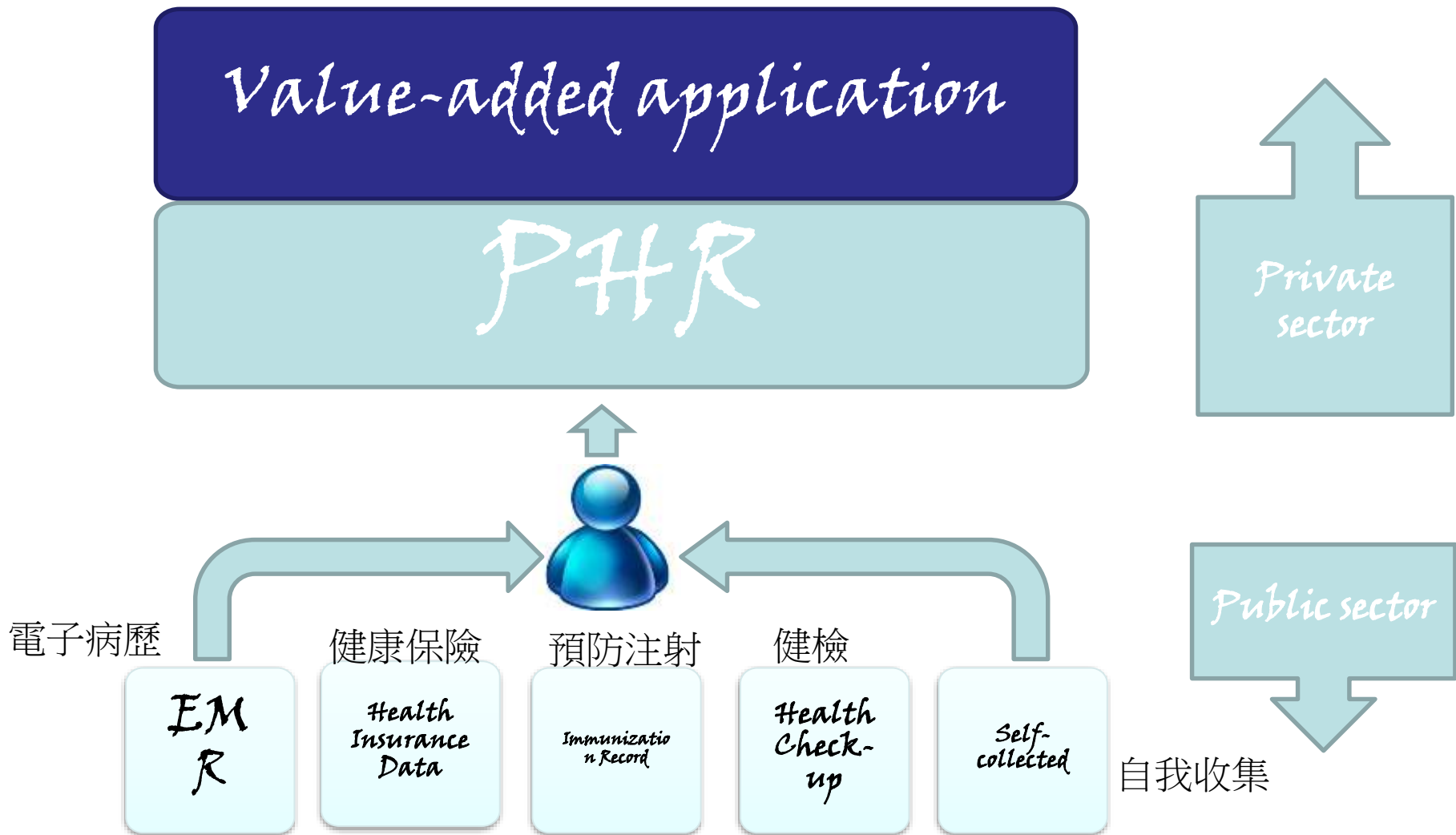- data are <u>secure</u> from loss, alteration or damage

衛生福利部推動新的應用與服務

- 健康雲端藥曆系統**(PharmaCloud)**，loud-based medication record system (PharmaCloud)
  - provide to query of patients medication
  - to avoid duplication of medication and drug adverse(interactions) events, overdose
- Launched in 2014 我的健康存摺**(My Health bank(Data) book)**
  - people can query their use of medical services in the past year, including hospitalization, medical expenses, outpatient clinic visits and prescription records, etc., people can learn more about their own health and develop better health lifestyle.
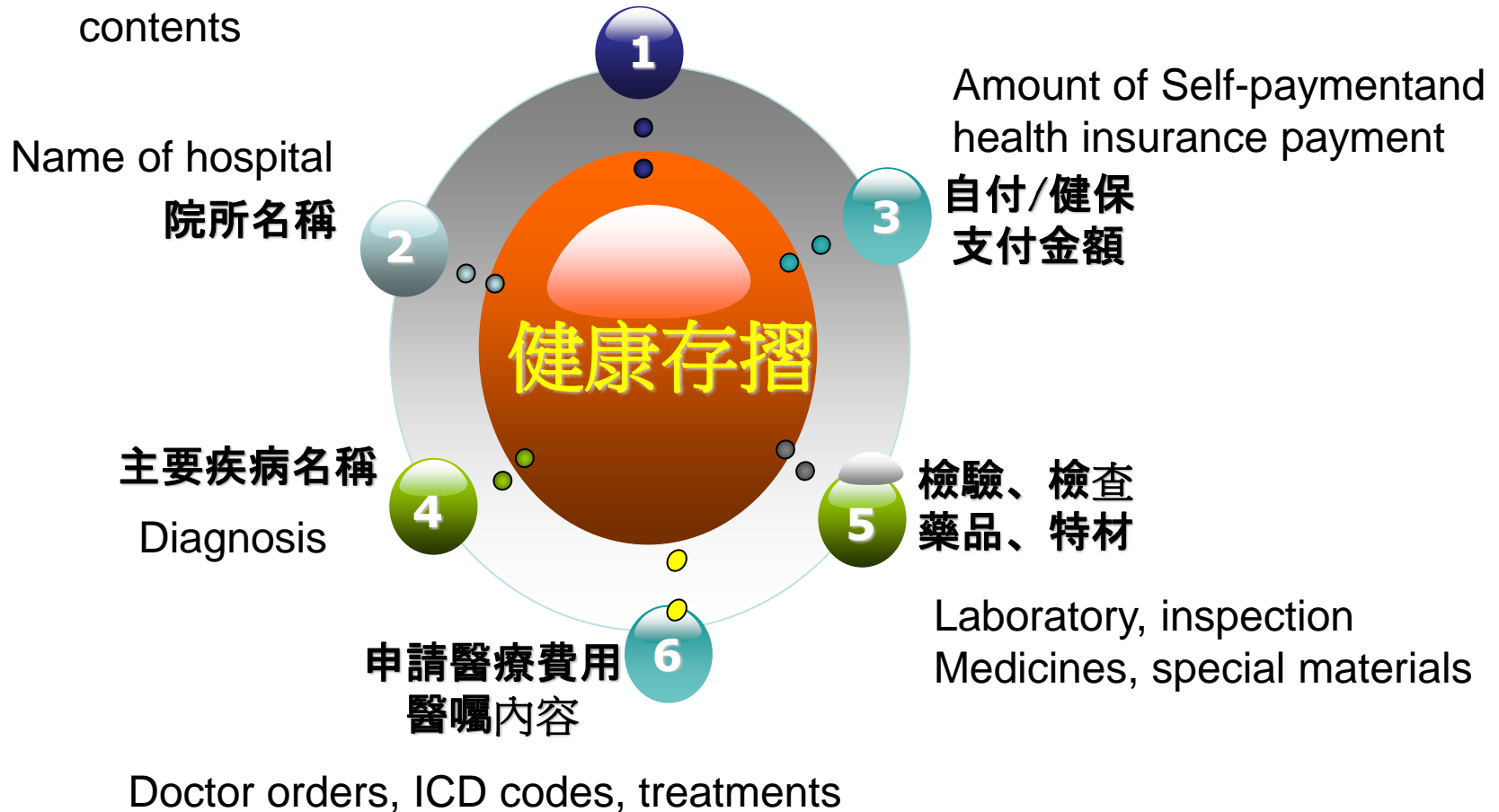
個人自我管理的電子健康紀錄
Personally controlled EHR

Value-added application

PHR

Private sector

Public sector

電子病歷
健康保險
預防注射
健檢

EMR

Health Insurance Data

Immunization Record

Health Check-up

Self-collected

自我收集

Min-Huei (Marc) Hsu, Director of Medical Informatics Center

全民健保健康存摺
My Health bank book

➤存放內容
contents

個人就醫或住院日期　　Dates of visiting of hospitals

Amount of Self-paymentand
health insurance payment
自付/健保
支付金額

Name of hospital
院所名稱

健康存摺

1

2

3

主要疾病名稱
Diagnosis

4

5

檢驗、檢查
藥品、特材

Laboratory, inspection
Medicines, special materials

申請醫療費用
醫囑內容

6

Doctor orders, ICD codes, treatments

32

Example of My Health Bank(Data) Book

<?xml version="1.0" encoding="utf-8"?>

<myhealthbank>

  <bdata>

   <b1.1>M12010****</b1.1>

   <b1.2>20180613</b1.2>

   <r1>

    <r1.1>1</r1.1>

    <r1.2>臺北業務組</r1.2>

    <r1.3>5931033489</r1.3>

    <r1.4>德美藥局</r1.4>　**Name of hospital**

    <r1.5></r1.5>

    <r1.6>20180527</r1.6>

    <r1.7>XXXX</r1.7>

    <r1.8>J00</r1.8>

    <r1.9>急性鼻咽炎（感冒）</r1.9>

    <r1.10></r1.10>　**ICD-10 Acute Nasopharyngitis (cold)**

    <r1.11></r1.11>

    <r1.12>0</r1.12>

    <r1.13>78</r1.13>

    <r1_1>

     <r1_1.1>A021841100</r1_1.1>

     <r1_1.2>祝敏康錠（匹普林奈）</r1_1.2>

     <r1_1.3>4.50</r1_1.3>

     <r1_1.4>3</r1_1.4>　**ATC code of drug**

    </r1_1>

    <r1_1>

     <r1_1.1>AC27528100</r1_1.1>

     <r1_1.2>滅嗽錠３０毫克（美蘇仿）</r1_1.2>

     <r1_1.3>9.00</r1_1.3>　**ATC code of drug and dose**

     <r1_1.4>3</r1_1.4>

    </r1_1>

    <r1_1>

     <r1_1.1>AC37603100</r1_1.1>

     <r1_1.2>"杏輝" 抑疼炎腸溶膜衣錠５０毫克（二克氯吩鈉）</r1_1.2>

     <r1_1.3>9.00</r1_1.3>　**ATC code of drug and dose**

     <r1_1.4>3</r1_1.4>

    </r1_1>

   </r1>

   <r1>

    <r1.1>1</r1.1>

    <r1.2>臺北業務組</r1.2>

    <r1.3>5931033489</r1.3>

    <r1.4>德美藥局</r1.4>

    <r1.5></r1.5>　**Name of hospital**

    <r1.6>20180603</r1.6>

    <r1.7>XXXX</r1.7>

    <r1.8>J00</r1.8>

    <r1.9>急性鼻咽炎（感冒）</r1.9>

    <r1.10></r1.10>

    <r1.11></r1.11>　**ICD-10 Acute Nasopharyngitis (cold)**

    <r1.12>0</r1.12>

    <r1.13>91</r1.13>

33

個人健康紀錄的一些應用

健康存摺應用案例

利用健康存摺增進病人用藥安全

Using My Health Bank Data Book to improve Patient Safety on checking the Potentially Inappropriate Medication Use in Older Adults

高齡不適當用藥警示

高齡須謹慎使用藥物提醒

高齡使用藥物之交互作用檢查

高齡依腎功能狀態之用藥建議

實現方法

Parsing the Data from My-HB

將健康存摺做 XML
解析**取出藥品代號**

Obtain the Pharmaceutical composition

將藥品代號對應到
健保用藥品項查詢
專案檔**取出成份**

健康存摺 2.0

Health Insurance
Drug Database

健保用藥
品項資料庫

Alert results returned

將 Beers Criteria 的
成份**警示結果回傳**

Beers Criteria
警示資料庫

將成份**比對 Beers
Criteria 警示資料庫**

健康存摺應用案例

利用健康存摺增進病人安全

利用健康存摺ICD診斷碼及 AirBox 空氣污染指標 提供呼吸系統疾病之空汙警示

Use health Bankbook (ICD diagnostic code) and AirBox Air Pollution Index to Provide warning of air pollution for respiratory disease patients

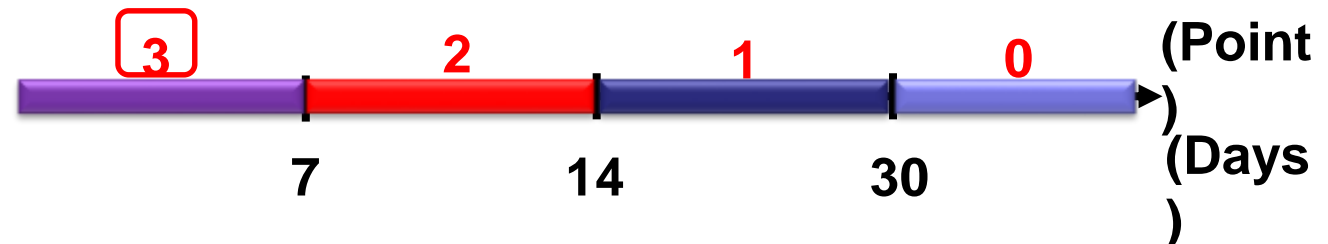# 空氣污染警示說明 Air Pollution Warning Depends on different Respiration Diseases
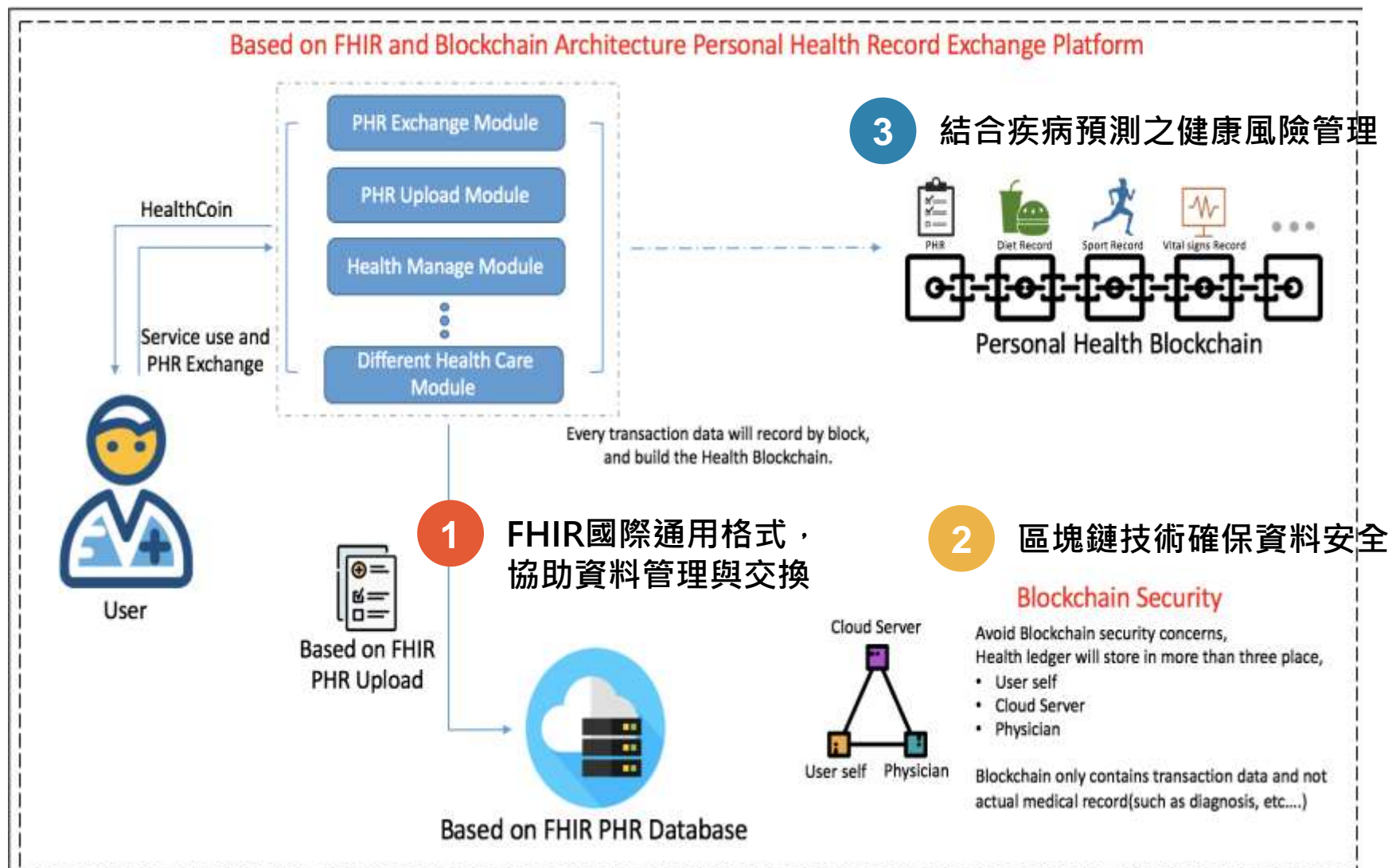
the Color of Air Pollution Map lighter is better

0 3 5 8 10 13 15 20 35 50 65 80 90 100

↓

Past 7 days had Bronchitis, acute(急性支氣管炎 ICD-9=466.0)

↓

Acute Respiratory Infections

| 3 | 2 | 1 | 0 | **(Point)** |
| 7 | 14 | 30 | | **(Days)** |
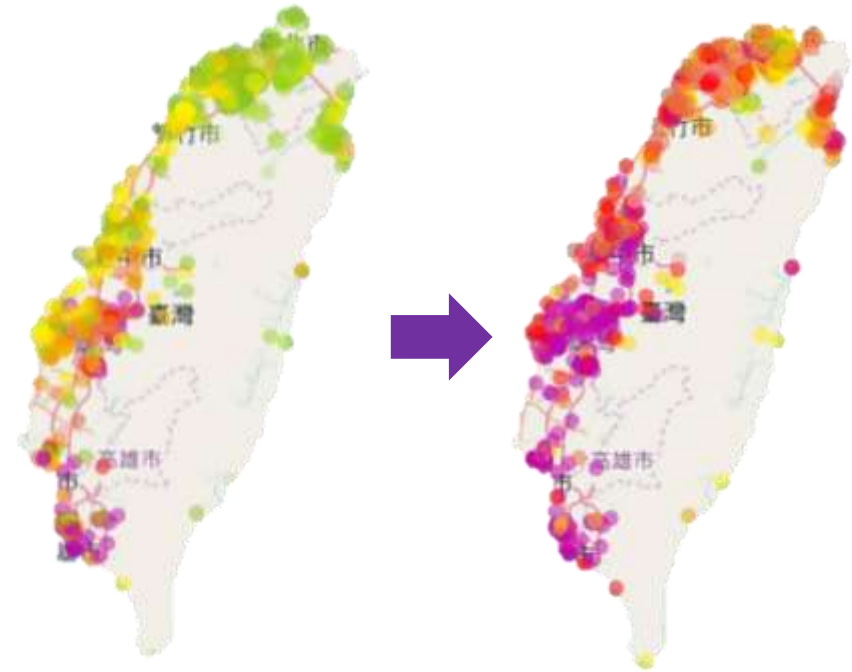
以FHIR及Blockchain 為基礎之
跨國際個人健康資訊紀錄交換平臺

Collaboration of Medical Information Sharing and Knowledge Exchange

- Taiwan: Standard and Interoperability Lab - Smart Healthcare center of Excellent Taiwan

- Asia eHealth Information Network (AeHIN): Standards and Interoperability Lab - Asia Community of Interoperable Labs (COIL)

- *University of the Philippines Manila*
- *Asian Development Bank*
- *National Taipei University of Nursing and Health Sciences*
- *Taiwan Electronic Health Association - TEHA*

40

醫學臨床數據庫的整合

The Value of use of health related data

- Secondary Use(加值應用或二次運用)去識別化之健康資料為世界趨勢
  - 美國早已在20年前開放全國住院資料供研究者使用, USA started from 20 years ago
- 新的治療方式、疾病的診斷、藥物之副作用、疾病之關聯性…等, for discovery of New treatment, diagnosis methods, side effects of drugs, relationship between diseases
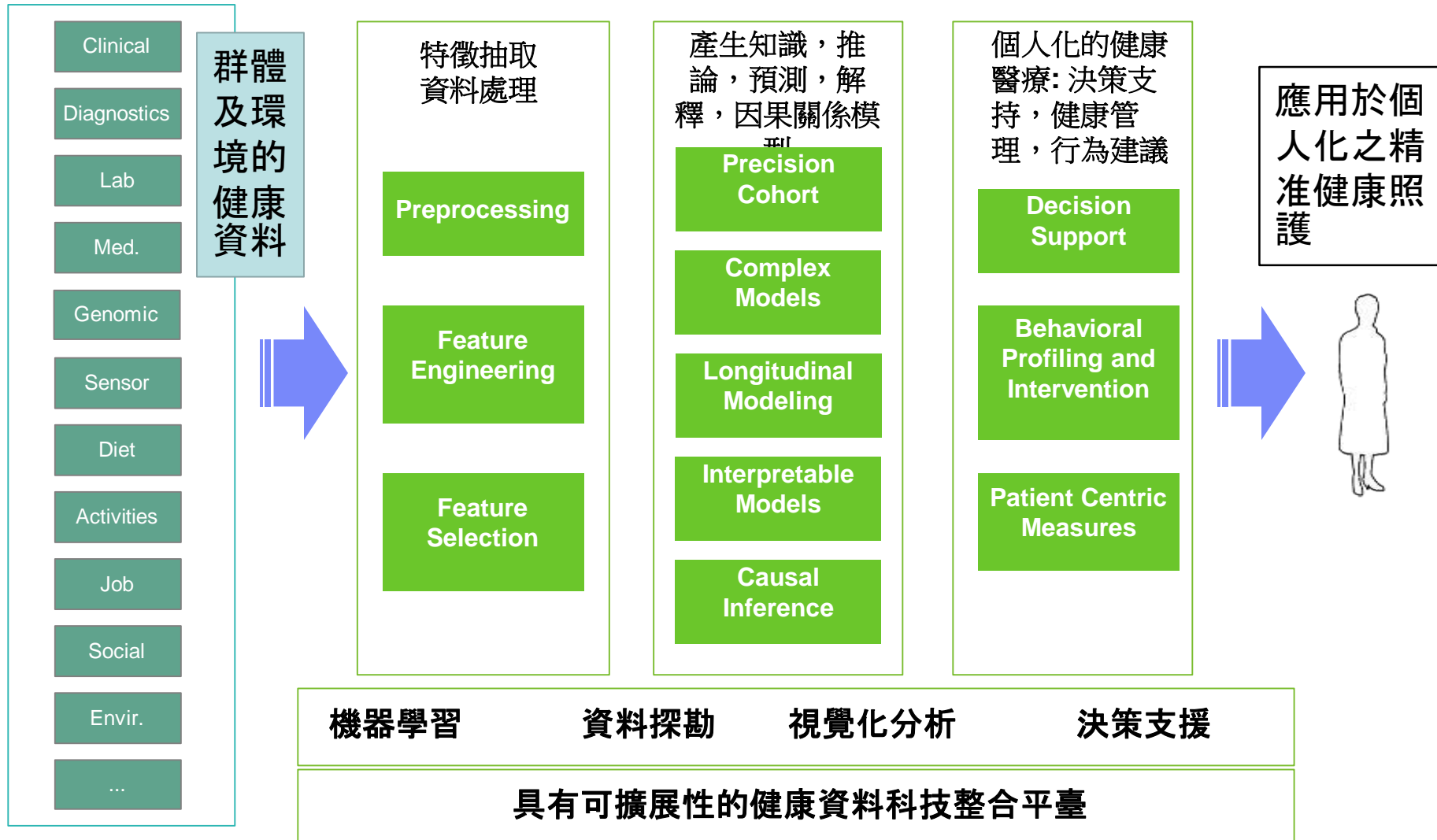- 若沒有完整開放健康加值資料將嚴重損害廣大病人之權益, open data will fit the benefits of majority of people

From Prof. Jack Li

Integrated Biomedical Informatics for Clinical Research
醫學研究之道，就是整合研究資料

- 2003年9月30日，美國國家衛生研究院（NIH）院長塞烏尼（Elias Zerhouni）宣佈：對美國政府資助的醫學研究進行重整。
- 「研究路線圖」的計畫 NIH Roadmap
- National Electronics Clinical Trials and Research (NECTAR) network
- 建立完整的路線圖以及徹底更新醫學資料的收集、儲存及共用 methods for collecting, storage, and sharing
- 把龐大且分散的資料庫結合成一個巨大的資料庫 integrating data bases
- 發展軟體，使實驗計畫的撰寫能夠簡化並標準化 developing software for helping experiment design and common data element
- 減少紙張的使用 reducing paper use

新世代健康照護的運算架構
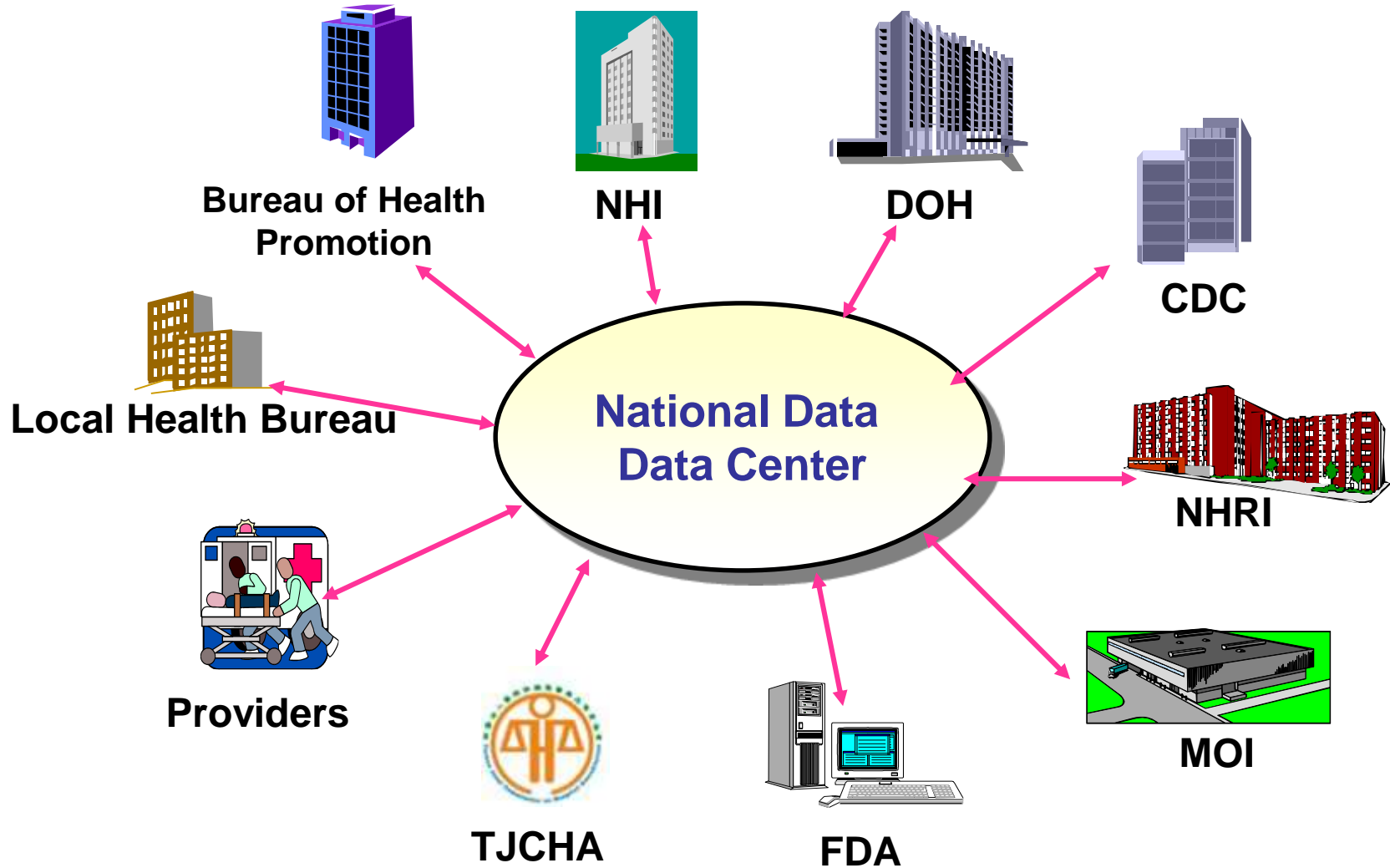# 利用群體的健康資料達到個人化的健康照護 from IBM Research

| Clinical |
| Diagnostics |
| Lab |
| Med. |
| Genomic |
| Sensor |
| Diet |
| Activities |
| Job |
| Social |
| Envir. |
| ... |

群體及環境的健康資料

特徵抽取資料處理

Preprocessing

Feature Engineering

Feature Selection

產生知識，推論，預測，解釋，因果關係模型

Precision Cohort

Complex Models

Longitudinal Modeling

Interpretable Models

Causal Inference

個人化的健康醫療: 決策支持，健康管理，行為建議

Decision Support

Behavioral Profiling and Intervention

Patient Centric Measures

應用於個人化之精准健康照護

機器學習　　　資料探勘　　　視覺化分析　　　決策支援

具有可擴展性的健康資料科技整合平臺

完整的健康醫學知識庫

- **Databanks related to health**

National Health Data Center

Health-related Databanks

Process and Outcome			
Mortality	**Morbidity**	**Disability**	**Others**
Death reporting	**1.Health insurance utilization databank** **2.Cancer registration**	**1.Catastrophic disease certificate databank** 2.**Disability certificate databank**	**1.Birth reporting** **2.Health care organization/service databank** **3.Special survey for care providers**
Health determinant			
Health behavior	**Economic status**	**Family status**	**Ethnicity**
Health and nutritional survey	**Household income and expenditure survey**	**Household & health insurance records**	**Indigenous people databank**

Collaboration Center for Value-added Application of Health Databanks – a Pilot (1)

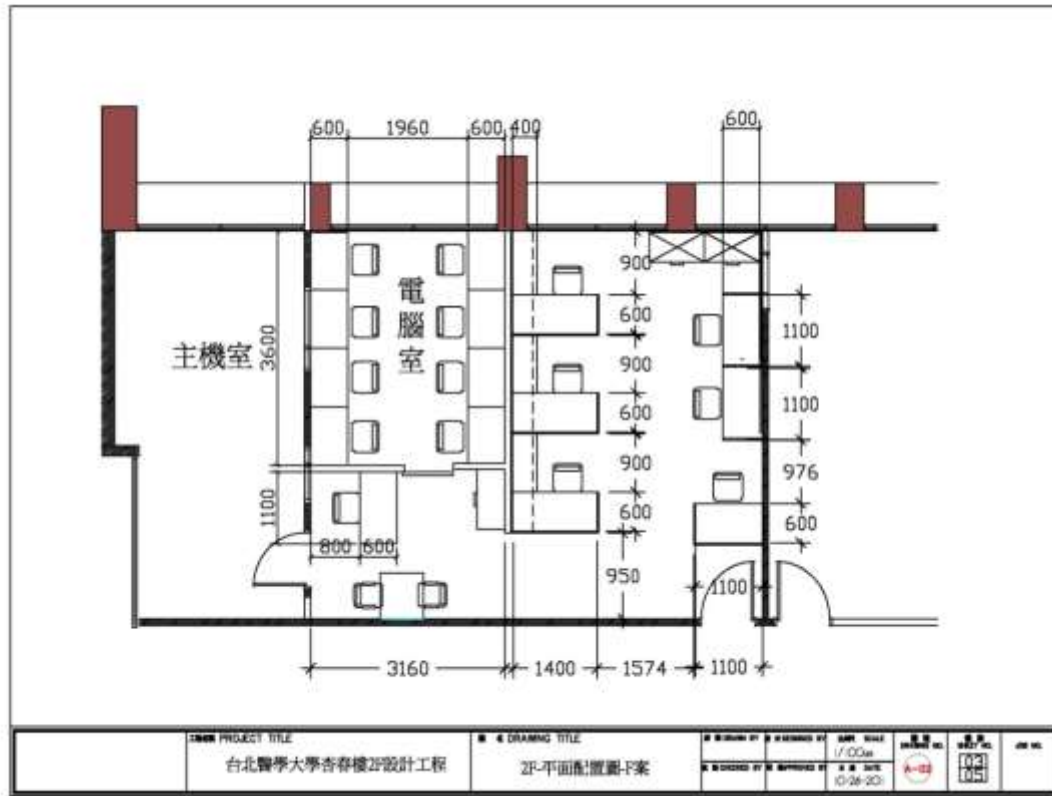***Sponsored by Taiwan DOH in collaboration with China Medical University (Taichung, Taiwan)**

臺北醫學大學健康暨臨床研究資料加值中心平面圖與現況圖Taipei Medical University

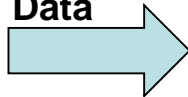左側隔間使用者坐位8個，2名管理人員坐位（左邊隔間左下角以及右邊隔間左下角）

主動式具有分析評估能力的主題式資料架構

Data Architecture by Subjects with Active Analysis and Assessment

Integrated Data Center for Bio-medical Informatics

Provide Data

Health Insurance DB

Cause of Death DB

Cancer Register DB

Household (census) Register DB
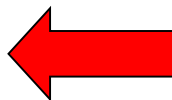
Redundant / Feedback

未通過專法，本中心不直接釋出資料

Limited Data Set

Continuity Research

Tool Kits

Regular Data Released

De-identification

Public Health Administration

Epidemiology

Health Care Management

:

Others

Data Released after IRB approval

User

Center of Health Information Application 2009

Scenario A

Integrated Data Center for Bio-medical Informatics

Health Insurance DB

Cause of Death DB

Cancer Register DB

Household (census) Register DB

Provide Data

1. 資料串接 Data joining
2. 匿名化處理 De-identification

資料處理與取用 Data processing and statistical Analysis according to query of the department of health
國民健康指標資料庫
National health indicator databank

現有資料庫種類(33種) different databases

代號	統計項目	資料提供單位
1	全民健保處方及治療明細檔-中醫門診、西醫門診、牙醫門診、住院、藥局 prescription details	健保局
2	全民健保處方及治療醫令明細檔-中醫門診、西醫門診、牙醫門診、住院、藥局	健保局
3	全民健保承保檔	健保局
4	全民健保重大傷病檔	健保局
5	戶籍資料檔 census file	內政部
6	死因檔 cause of death	統計室
7	醫事機構現況檔	統計室
8	醫事機構服務量檔	統計室
9	醫療院所評鑑等級	醫事處
10	出生通報檔	國民健康局
11	癌症登記檔(含癌症登記檔、癌症診療資料庫、短表、長表)	國民健康局
12	家庭醫療保健消費調查資料檔	統計室
13	醫事機構基本檔 medical institutes	健保局
14	原住民檔(區分出生、死亡、戶籍)	原住民委員會

Precision Medicine Moves Care from Population-Based Protocols to Truly Individualized Medicine

President Obama announced the Precision Medicine Initiative in his 2015 State of the Union address. (精準醫學倡議)

The Exposome: Exposure to Disease

Our Exposures

Our Unique Characteristics

Disease

- Under the initiative, medical care would transition from a one-size-fits-all approach to an individualized approach, in which data on each patient's genomic makeup, environment, and lifestyle (基因組組成，環境和生活方式, the exposome, Exposome and Exposomics) helps medical professionals tailor treatment and prevention strategies.

- To achieve the Precision Medicine Initiative mission statement, "to enable a new era of medicine through research, technology, and policies that empower patients, researchers, and providers to work together toward development of individualized care," researchers and clinicians need vast and varied amounts of data and the technology to ensure that data is widely accessible and usable. (資料和技術，以確保資料可廣泛使用)

A.I./Machine Learning Methodology for Biomedical Datasets
數據共用下的機器學習的步驟

"High Quality" &
large quantity of
Biomedical
Dataset(s)

大量高品質的資料

生物和臨床驗證

Biological
& Clinical Validation
Actionable A.I.

Features Selection
Label/Annotation
(Domain Knowledge)

**Data
Commons**
資料共用

特徵選擇及標注(需
要專業領域知識)

統計和生物醫學解釋

Statistical &
Biomedical
Interpretation

Machine Learning
Neural Network
Pipeline/
Optimization

新的資料分析技術
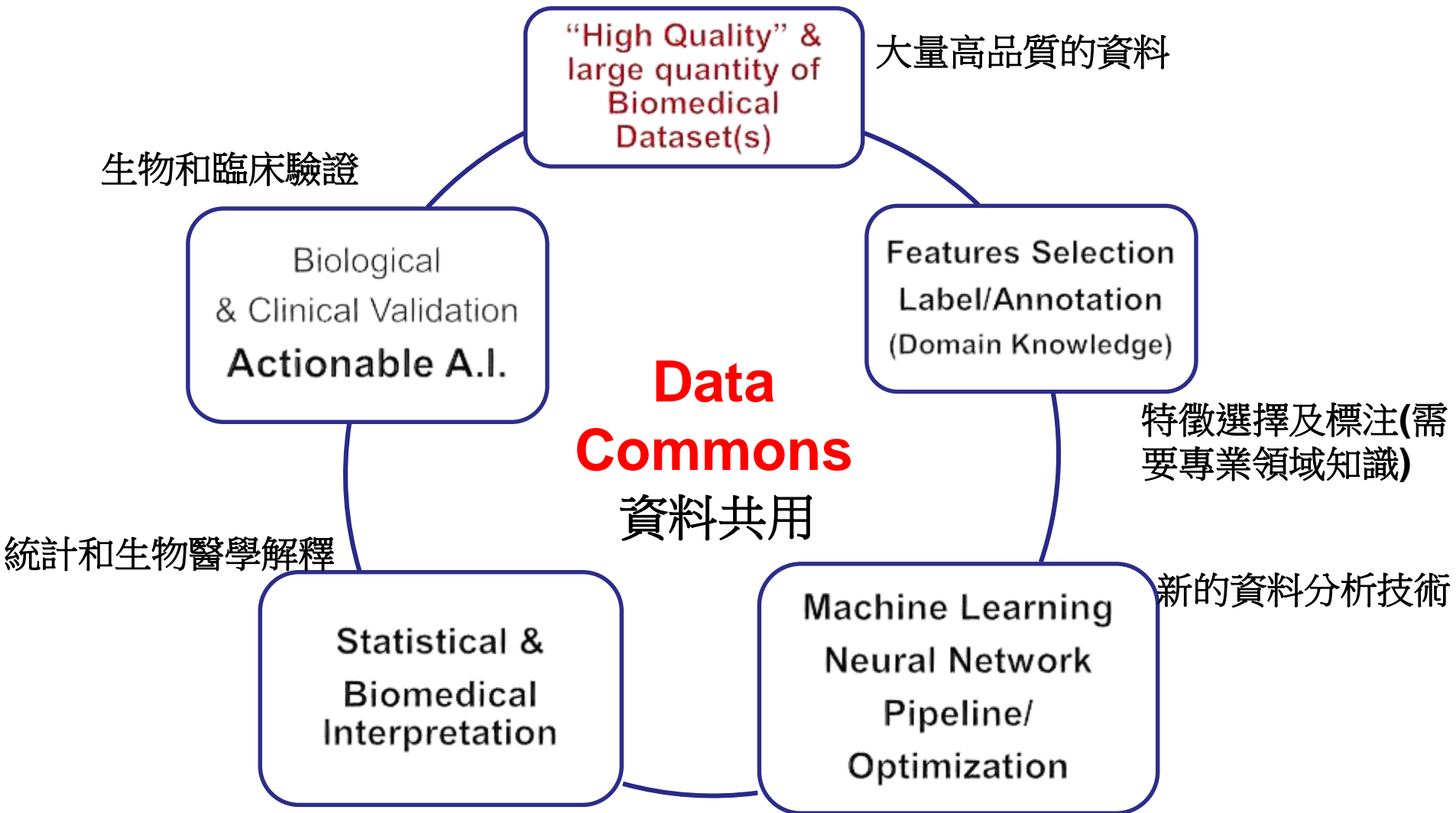
電子病歷系統

國家級人體生物資料庫整合平台
建置標準醫療資訊模式--台灣癌症通用資料庫

基本評估

Lab & Vital signs

病人

處置&藥物

Out-come
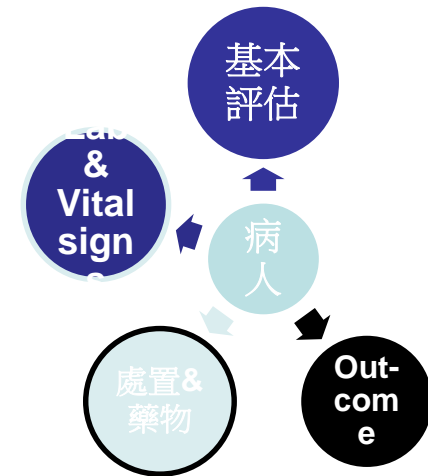
Common Data Model (CDM)
XML檔案格式

- 使用 R studio 處理 **data management & analyses**
- 平台辦公室提供 **centralized analysis programs**

- **cancer-specific database**
- **project-specific database**
- 各項主題資料庫

Biobank
資料庫申請者

醫療機構的**biobank**

電子病歷系統

基本評估

Lab & Vital sign

病人

處置&藥物

Out-come

Common Data Model (CDM)
XML檔案格式

● 使用 **data & a** ● 平台 **cen** **ana**

● **car dat** ● **pro dat** ● 各項

醫療機構的biobank

● 第一階段工作,先彙整既有的結構式欄位,
● 主要針對癌症治療成效分析所需要的欄位,
● 以符合查驗登記**(臨床試驗)**所需的資料欄位為基礎

資料擷取與應用流程
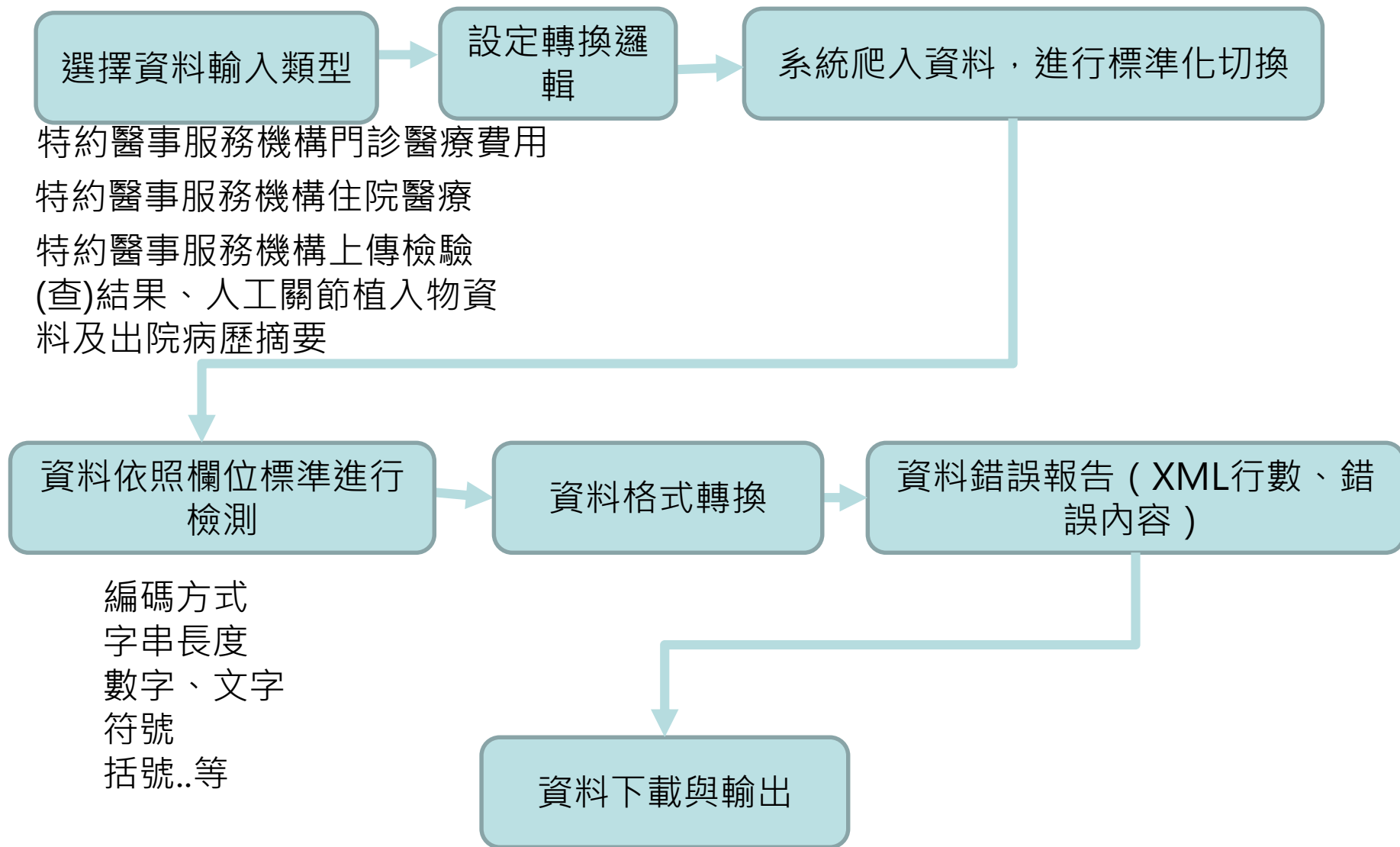
選擇資料輸入類型 → 設定轉換邏輯 → 系統爬入資料，進行標準化切換

特約醫事服務機構門診醫療費用

特約醫事服務機構住院醫療

特約醫事服務機構上傳檢驗
(查)結果、人工關節植入物資
料及出院病歷摘要

資料依照欄位標準進行檢測 → 資料格式轉換 → 資料錯誤報告（XML行數、錯誤內容）

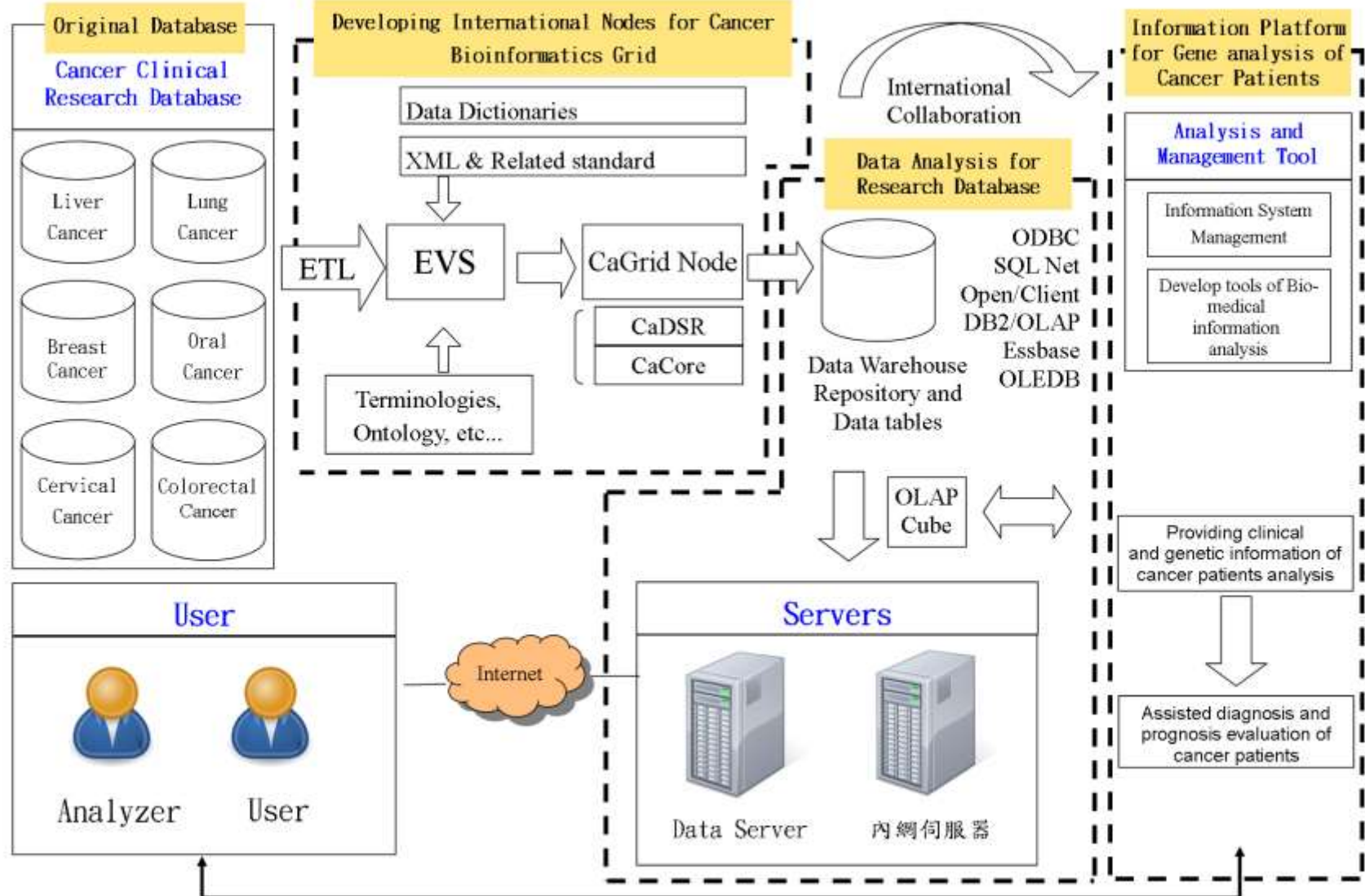編碼方式
字串長度
數字、文字
符號
括號..等

資料下載與輸出

Taiwan Information Integrated Platform for Cancer Research : cancer Data Standards Repository and Cloud Computing

資料的建立　　統一的格式及國際標準　　資料分類整合及應用　分析工具及平台

Real-World Data Implementation

RWD 用於我國醫療科技評估的現況

- 目前我國醫療科技評估評估案件可分為三大類，
  1. 新醫療科技(新藥/新醫療器材/ 醫療服務)建議收載前評估
  2. 健保署委託之特定主題研究案
  3. 衛福部之醫療衛生介入政策之研究案

- 台灣的全民健保的涵蓋率高達 99%，健保資料庫及癌症登記 檔等大型資料庫的數據資料相對完善，亦是台灣最常使用且最完整的 RWD 之一。我國醫療科技評估已將健保資料庫，廣泛運用在估算疾病盛行率、疾病發生率、資源利用、 醫療成本，及經濟評估模型之轉移機率(transition probability)等。

- 然而，RWD 在相對療效方面的運用則相對較少，主要原因可歸納出以下幾點:
  - 健保資料庫並無涵蓋尚未收載給付的藥品，故建議收載前，無法利用健保資料庫進行相對療效評估
  - 健保資料庫無記載實驗室檢驗數據，在療效評估有所限制
  - 建議收載前執行醫療科技評估的作業天數太短，無法額外進行RWD蒐集
  - 決策者對於RWD 研究不熟悉
  - 醫療決策者對 RWD 研究所產生的 RWE 仍具不確定性

Real world data 介紹及於醫療科技之應用
https://www.cde.org.tw/Content/Files/Knowledge/90073b31-4404-44db-80aa-13a68173ffb5.pdf

國際藥廠運用 Real World Data 藥品上市案例(Ibrance)

- **How real world evidence was used to support approval of Ibrance for male breast cancer (The Cancer Letter, 20190419, https://cancerletter.com/articles/20190419_2/)**
  - 2015 NDA approval for female breast cancer
  - 2016-2017 US FDA regulatory meeting
  - 2018 submit to US FDA for use indication
  - 20190419 US FDA approval by RWE
- FDA said:
  - Pfizer provided the results of an analysis of real world data (RWD) from electronic health records (EHRs) as additional supportive data to characterize the use of Ibrance in combination with endocrine therapy (aromatase inhibitor or fulvestrant) in male patients with breast cancer based on observed tumor responses in this rare subset of patients with breast cancer.
  - Leveraging RWD to improve regulatory decisions is a key strategic priority for the FDA. This data may be derived from a variety of sources, such as electronic health records, medical claims, product and disease registries, laboratory test results and even cutting-edge technology paired with mobile devices.
  - These types of data are being used to develop real world evidence (RWE) that can better inform regulatory decisions(監管決策提供真實的證據依據).
  - RWD has been used to determine prognosis or natural history of disease in order to help inform regulatory decision-making, for example, data on historical response rates drawn from expanded access(擴展訪問), practice settings, or chart reviews.
  - Real world tumor response data: electronic health record and information about response event collected retrospectively. physical exam, symptom improvement, pathology reports, and durations of prescriptions.
  - The safety and efficacy of Ibrance had been previously demonstrated in women and therefore was only supported by RWD in order to extend the indication to men. (*primary evidence in this case*)

國際藥廠運用 Real World Data 藥品上市案例(Ibrance)

- Male breast cancer
  - less than 1% of all cases of breast cancer
  - not feasible to conduct prospective clinical studies
- Using Real World Data to get RWE
  - post-marketing reports
  - electronic health records databases
    - IQVIA Insurance database (健康保險資料)
    - Flatiron Health Breast Cancer database: baseline characteristics, treatment patterns and clinical outcomes from patient-level, de-identified data for a group of male patients with metastatic breast cancer. (the data is "fit-for-purpose," that is, ensuring that the dataset is fit for the intended use and can provide adequate scientific evidence) (電子病歷，BioBanks and other臨床試驗 data)
    - the Pfizer global safety database (藥物安全資料庫)
  - Based on these data, the safety profile for men treated with IBRANCE is consistent with the safety profile in women treated with IBRANCE.
- Real world data are playing an increasingly important role in expanding the use of already-approved innovative medicines
- The 21st Century Cures Act, enacted in 2016, was created to help accelerate medical product development, allowing new innovations and advances to become available to patients who need them faster and more efficiently. This law places additional focus on the use of real world data to support regulatory decision-making.
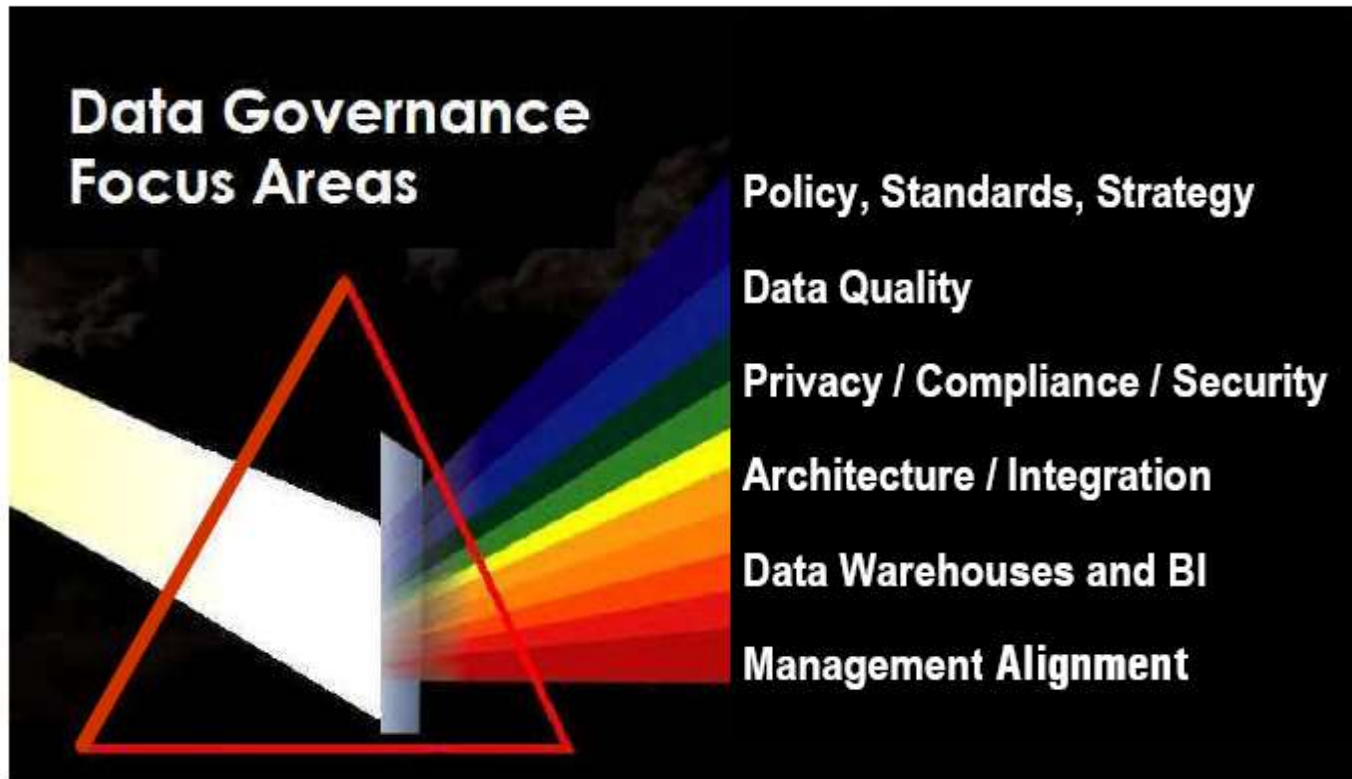
Want can Data Governance Accomplish?資料治理的重要

- Enable better decision-making
- Reduce operational friction
- Protect the needs of data stakeholders
- Train management and staff to adopt common approaches to data issues
- Build standard, repeatable processes
- Reduce costs and increase effectiveness through coordination of efforts
- Ensure transparency of processes

The most common objective of Data Governance programs is to standardize data definitions across an enterprise or initiative. (for example: 電子病歷標準)

Typical Data Governance Focus Areas

Seven phases of Data Governance Life Cycle Methodology

智慧化以個人為中心之健康照護資訊應用架構
Personal Health Information Application Scenario NHIP

Diet Record

Exercise Record

BP, BS Record

Others

Public Health Knowledge Portal

Health Examination Center

Health Examination Center

People

Hospital

Hospital

Health Management Institution

Health Management Institution

Health Information Exchange Platform

Nursing Care Institution

Health Information Exchange Platform

Personal Health Data Bank

Health Information Exchange Platform

Nursing Care Institution

Q&A

Prof. Chien-Yeh Hsu 徐建業教授 0939193212
e-mail:cyhsu@ntunhs.edu.tw
臺北護理健康大學
臺北醫學大學
台灣醫學資訊學會
臺灣護理資訊學會
臺灣電子健康學會
National Taipei University of Nursing and Health Sciences
Taipei Medical University
Taiwan Nursing Informatics Association, TNIA
Taiwan Association for Medical Informatics, TAMI
Taiwan Electronic Health Association - TEHA