

Clinical Data Harmonization for AI in Healthcare

Prapat Suriyaphol

Faculty of Medicine Siriraj Hospital

3 Oct 2023

พันธกิจ SiData+

Data Innovation

1

Data Governance and
Regulation

3

Data Consultation and
Internal Collaboration
on Data

4

Research and
Development for
Data Utilization

5

External
Collaboration on
Data

2

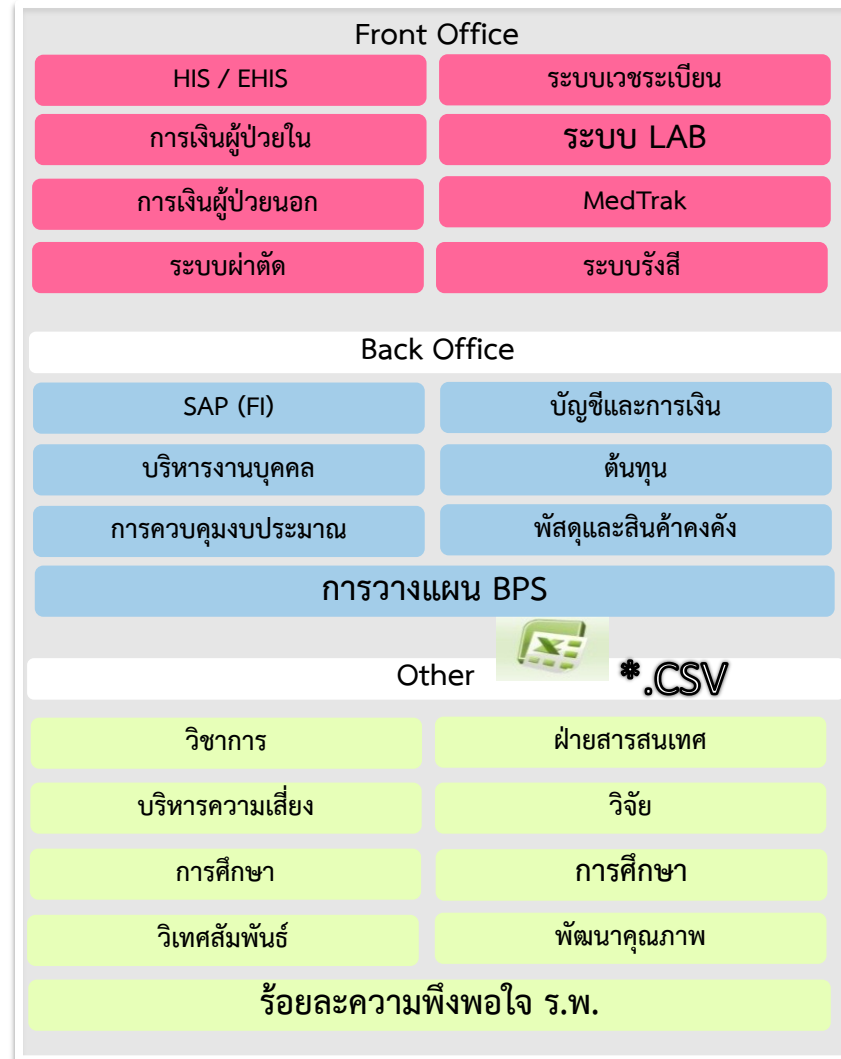
Data Management

1. กำกับธรรมาภิบาลข้อมูล ดูแลและควบคุมการใช้งานข้อมูล (Data Governance and Regulation) ตาม พ.ร.บ. การบริหารงานและการให้บริการภาครัฐผ่านระบบดิจิทัล พ.ศ. 2562 และ พ.ร.บ.คุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562
2. รวบรวมและบริหารจัดการข้อมูล (Data Management)
3. ให้คำปรึกษาและร่วมมือกับหน่วยงานต่าง ๆ ภายในคณะแพทยศาสตร์ศิริราชพยาบาล (Data Consultation & Internal Collaboration on Data)
4. วิจัยและพัฒนาการนำข้อมูลมาใช้ประโยชน์ (Research and Development for Data Utilization) ด้วยหลักการวิเคราะห์ข้อมูล (Data Analytics) และวิทยาศาสตร์ข้อมูล (Data Science)
5. อำนวยความสะดวก กำกับดูแล และควบคุมการใช้งานข้อมูลของคณะแพทยศาสตร์ศิริราชพยาบาลโดยหน่วยงานภายนอก (External Collaboration on Data)

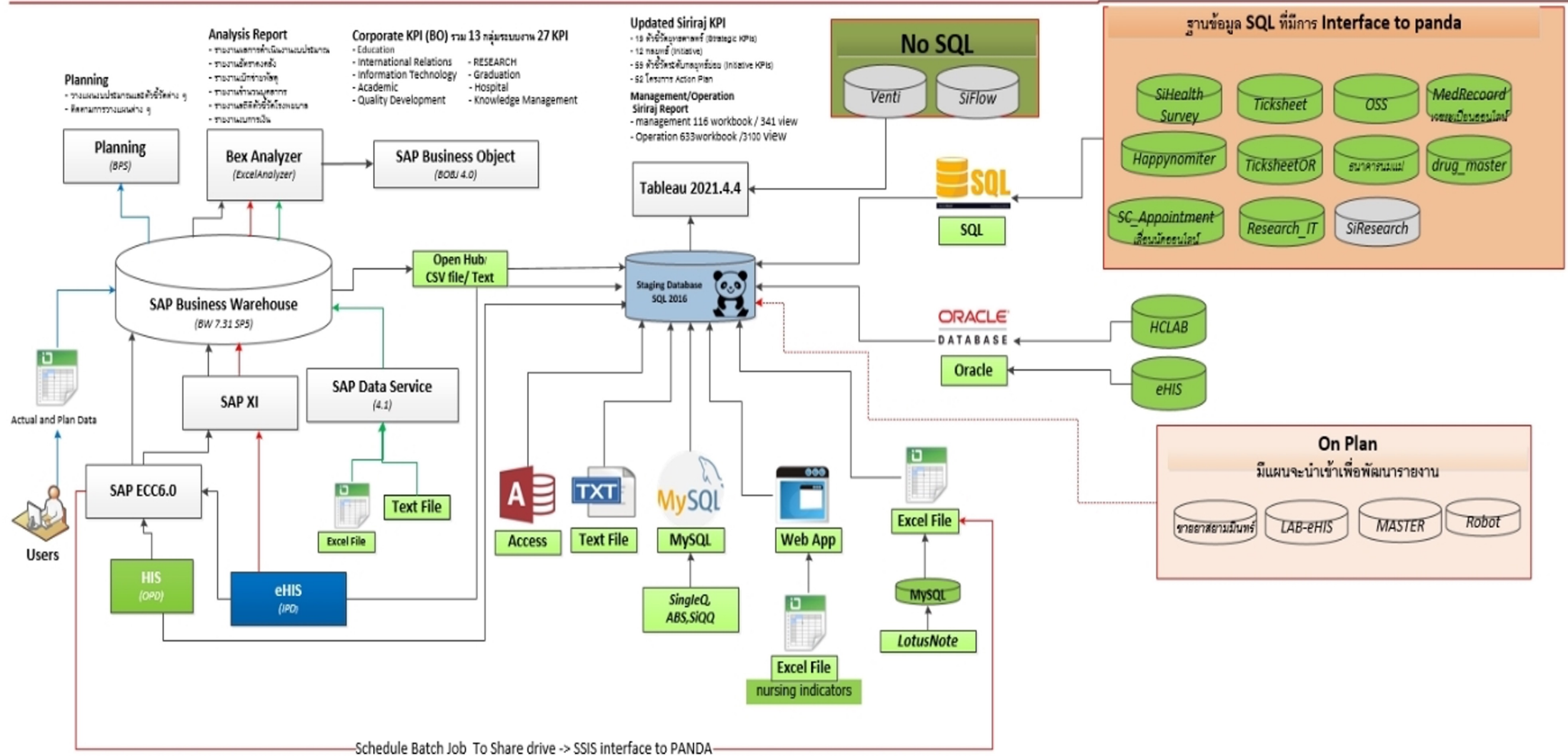
ทีม Data Governance



DATA SOURCE



Current Siriraj's BI + Tableau Architecture



Accessible Data at Siriraj

Data Type	Description
Patient Electronic Health Records (EHR)	Patient medical records, including medical history, lab results, and other health information
Medical Imaging Data	Data from medical imaging tests, such as X-rays, CT scans, and MRIs, Slide scan from Pathology depart.
Genomic Data	Data from genetic tests and sequencing
Drug and Medication Databases	Databases of drugs and medications, including information on side effects and interactions
Telehealth Data	Data from telehealth visits, such as video calls with doctors
Health Insurance Claims Data	Data from health insurance claims, including information on treatments and costs
Electronic Prescription Data	Data from electronic prescriptions, including information on medications prescribed
Hospital and Clinic Operational Data	Data from hospital and clinic operations, such as patient wait times and staff schedules

Siriraj Informatics and
Data Innovation Center



Mahidol University
Faculty of Medicine
Siriraj Hospital

Hands-on Workshop: Introduction to International Medical Data Standard OMOP CDM for Observational Research

21 September 2023

Supported by



ศูนย์ความเป็นเลิศด้านชีววิทยาศาสตร์ (องค์การมหาชน)
Thailand Center of Excellence for Life Sciences
(Public Organization)





Function
Unique
representation of a
source code

Function
Used for standardized
analytics and by
OHDSI tools

condition_source_value	condition_source_concept_id	condition_concept_id	Standard Code
I10 <i>(ICD10 for Essential (primary) Hypertension)</i>	45591453	320128	59621000 <i>(SNOMED for Essential hypertension)</i>

การใช้งาน OMOP CDM จะต้องแปลงรหัสการบันทึกข้อมูลให้เป็นไปตามมาตรฐานที่ตกลงไว้ เช่น

รายการ	รหัสที่ใช้ในไทย	รหัสมาตรฐาน
โรค	ICD 10 TM	SNOMED-CT
หัตถการ	ICD 9 CM	ICD 9 Proc, SNOMED-CT
ยา	TMT	RxNorm
การตรวจทางห้องปฏิบัติการ (Lab)	TMLT	LOINC

Siriraj Informatics and
Data Innovation Center



Mahidol University
Faculty of Medicine
Siriraj Hospital

Data Governance for Research

Supported by



ศูนย์ความเป็นเลิศด้านชีววิทยาศาสตร์ (องค์การมหาชน)
Thailand Center of Excellence for Life Sciences
(Public Organization)



Data Governance

2 Prongs

Promotion

*Support and Enhance
Data Uses*

What are the data we have?

Where are the data?

How are the data collected and used?

Whose data is it?

How can we improve data quality?

How can we make the data more valuable?



Deliverables

Policies, Procedures,
Data Catalog,
Metadata,
Data Lineage,
Data Quality Assurance

Regulation

*Govern and Take Care of
Data Uses*

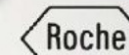
How can we use the data given legal regulations & ethics guidelines?

How can we facilitate external parties' usage of our data assets given intellectual property & legal considerations?



Secondary Use of Health Data

European review by Open Data Institute





โครงการพัฒนาระบบเชื่อมโยงข้อมูลสุขภาพของผู้ป่วย เพื่อยกระดับความมั่นคงสุขภาพในการพัฒนาระบบบริการและนโยบายด้านสุขภาพ: การเชื่อมโยงข้อมูลบริการสุขภาพผู้สูงอายุของโรงพยาบาลศิริราชและสำนักงานหลักประกันสุขภาพแห่งชาติ

โครงการนี้ได้รับทุนสนับสนุนจาก ศูนย์ความเป็นเลิศ
ด้านชีววิทยาศาสตร์ (องค์การมหาชน)
ปีงบประมาณ 2566 สัญญาเลขที่ TC (ERP) 13/2566



ที่มาและความสำคัญ

การวิจัยทางการแพทย์ เพื่อตอบสนอง โจทย์ต่าง ๆ
รวมถึง โรคระบาดระดับชาติและโรคอุบัติใหม่
มีประสิทธิภาพมากขึ้นจากการใช้ข้อมูลจากหลายระบบร่วมกัน



ข้อมูลจากระบบโรงพยาบาล (EHR) และระบบการเบิกจ่าย
(Claims) มีโครงสร้างแตกต่างกัน ใช้งานร่วมกันได้ยาก



การปรับโครงสร้างข้อมูล ให้ใช้มาตรฐานร่วมเดียวกัน
(OMOP CDM) ทำให้เชื่อมโยงข้อมูล (Data Linkage) ง่ายขึ้น



ผลการดำเนินโครงการ

ต้นแบบระบบการเชื่อมโยงข้อมูลจากฐานข้อมูลสุขภาพของ โรงพยาบาล
ศิริราชและฐานข้อมูลสำนักงานหลักประกันสุขภาพแห่งชาติ รวมถึงระบบ
การพัฒนาฐานข้อมูลต้นแบบที่สามารถขยายการเชื่อมโยงกับฐานข้อมูล
อื่นในอนาคตได้

- เครื่องมือการเตรียมระบบ โครงสร้างพื้นฐาน (infrastructure)
- เครื่องมือการปรับข้อมูลตามมาตรฐานนานาชาติ OMOP CDM
- เครื่องมือการเชื่อมโยงข้อมูล (Data Linkage)
- คู่มือการใช้เครื่องมือในระบบฯ สำหรับนักวิจัย
- คู่มือการปิดบังข้อมูลระดับบุคคล (De-identification)
- ร่างนโยบายธรรมาภิบาลข้อมูล



หัวข้องานวิจัยที่มีความสำคัญและสามารถนำไปใช้ประโยชน์ในการ
พัฒนานโยบายต่อไปได้

- ภาวะเปราะบาง (Frailty) ในผู้สูงอายุ
- COVID-19 ในผู้สูงอายุ
- ภาวะกระดูกสะโพกหัก (Hip Fracture) ในผู้สูงอายุ



จัดอบรมการใช้งาน OMOP
CDM และเครื่องมือภายใน
ระบบเชื่อมโยงข้อมูลในการ
วิจัย



ความร่วมมือระหว่างเครือข่ายนัก
วิจัย และผู้กำหนดนโยบายของ
สปสช. ด้านการวิจัยและพัฒนา
องค์ความรู้ร่วมกัน



แผนการดำเนินการต่อเพื่อขยายผลลัพธ์

ขยายการใช้งานระบบต้นแบบ ให้นำไปสู่การพัฒนาผลงานวิจัยจากการ
เชื่อมโยงข้อมูลหลากหลายแหล่ง ทั้งในประเทศและต่างประเทศ

- ปรับข้อมูลตามมาตรฐาน OMOP CDM ใน โรงพยาบาลศิริราช และ สปสช. ให้ครอบคลุมข้อมูลมากยิ่งขึ้น
- เพิ่มเติมและตรวจสอบการจับคู่ (map) รหัสโรค (ICD-10-TM) หัตถการ (ICD-9-CM) ยา (TMT) และการตรวจทางห้องปฏิบัติการ (TMLT) ให้เป็นไปตามมาตรฐาน OMOP CDM ได้แก่ SNOMED, ICD9Proc, RxNorm และ LOINC ตามลำดับ
- ขยายการปรับข้อมูลตามมาตรฐาน OMOP CDM ไปยัง โรงพยาบาลที่ สนใจเข้าร่วม

พัฒนางานวิจัย ที่มีความสำคัญและสามารถนำไปใช้ประโยชน์ใน
การพัฒนานโยบายต่อไปได้ ได้แก่ 3 หัวข้อข้างต้น และเพิ่มเติม
หัวข้ออื่น ๆ โดยอาจเข้าร่วมกับเครือข่ายวิจัยระดับนานาชาติ ซึ่ง
เป็นการเตรียมความพร้อมนักวิจัย เพื่อตอบสนอง โจทย์ต่าง ๆ รวมถึง
โรคระบาดระดับชาติและโรคอุบัติใหม่ ในอนาคตต่อไป

ขยายการจัดอบรม เพิ่มเติมหัวข้อ
การปรับข้อมูลตามมาตรฐาน การ
วิเคราะห์ข้อมูลระดับสูงเพื่อการ
วิจัย รวมถึงจัดทำคู่มือการใช้งาน
OMOP CDM ภาษาไทยที่เข้าถึง
ได้ที่ <https://omop.sidata.plus>

ขยายความร่วมมือ สร้างเครือ
ข่ายนักวิจัยและสถาบันต่าง ๆ
เพื่อให้เกิดการปรับข้อมูลตาม
มาตรฐาน การเชื่อมโยงข้อมูล
และการนำข้อมูลไปใช้เพื่อการ
วิจัย อย่างต่อเนื่องและติดตาม
ผลลัพธ์

ต้นแบบระบบการเชื่อมโยงข้อมูลจากฐานข้อมูลสุขภาพของโรงพยาบาลศิริราช และฐานข้อมูลสำนักงานหลักประกันสุขภาพแห่งชาติ รวมถึงฐานข้อมูลต้นแบบที่สามารถขยายการเชื่อมโยงกับฐานข้อมูลอื่นในอนาคต

ลำดับ	รายการ	Code Repository	หมายเหตุ
1	System Core Modules	https://github.com/sidataplus/omop-core	มีคำอธิบาย Architecture ประกอบ
2	OMOP Vocabulary Loader	https://github.com/sidataplus/omop-vocab-loader	มีคู่มือประกอบ
3	OMOP Data Quality Dashboard	(ใช้ code จาก https://github.com/OHDSI/DataQualityDashboard)	
4	OMOP Data De-identification	(ใช้ code จาก https://github.com/OHDSI/CureIdRegistry)	มีคู่มือประกอบ
5	OMOP Data Linkage	https://github.com/sidataplus/omop_linkage_dbt	มีคู่มือประกอบ
6	OMOP Atlas Deployment	(ใช้ code จาก https://github.com/OHDSI/Broadsea และ https://github.com/microsoft/OHDSIonAzure)	มีคู่มือประกอบ แบ่งเป็น (1) Broadsea ด้วย Docker (2) OHDSI on Azure สามารถเข้าถึงได้ที่ https://atlas-demo.sidata.plus/atlas
7	OMOP Atlas Manual	(ใช้ code จาก https://github.com/OHDSI/Atlas และ https://github.com/OHDSI/WebAPI)	มีคู่มือประกอบ



OHDSI Thailand Chapter?

