A success story

The Adoption of the OMOP-CDM in South Korea

AJOU UNIVERSITY | OHDSI

Rae Woong Park MD, PhD

Ajou University School of Medicine South Korea

Disclosures

Founder and stockholder of
EvidNet   evidnet

- OMOP CDM-based distributed research network platform startup company in Korea

- One-step CDM conversion service: legacy EMR data to OMOP CDM

- Vocabulary mapping service: mapping of local vocabulary to OMOP standard vocabulary

- Maintenance service: routine maintenance of CDM DB (ex) CDM version/vocabulary updates, integration of new CDM extensions, etc.

- Distributed research network platform for participating hospitals (in development)

Contents

About South Korea
Selected Indicators for Korea

GDP (nominal) 2017 estimate

Total
$1,498 trillion
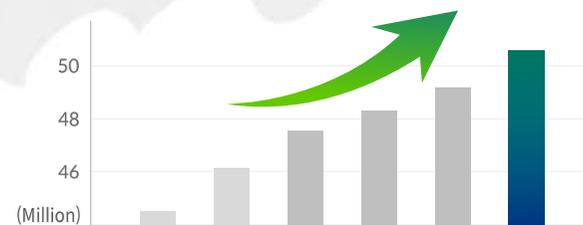(11th)

Per capita
$29,114
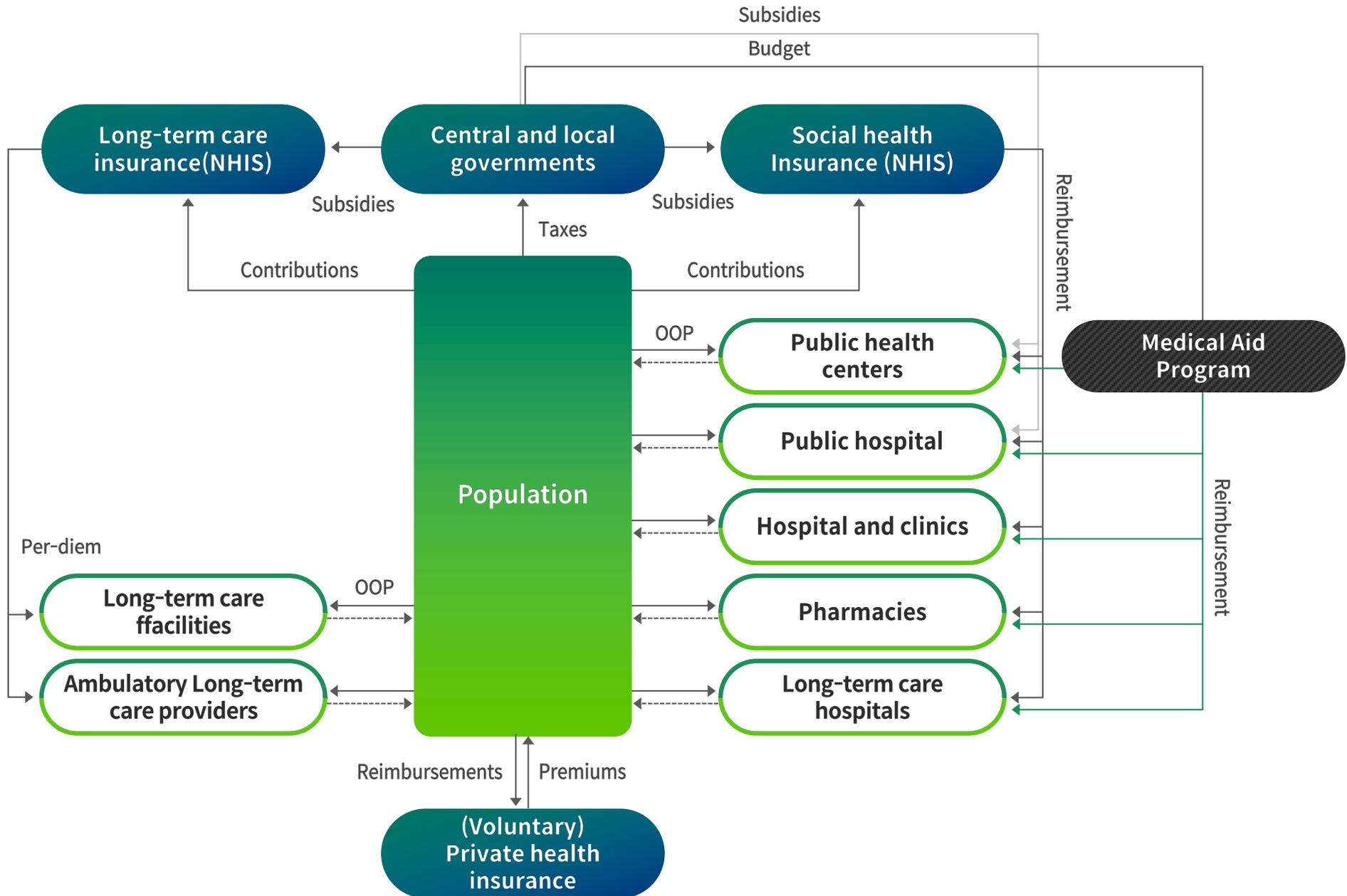(27th)

Life expectancy

Men
77 years

Women
84 years

Population

50
48
46
(Million)

50.4 million
Growth rate : 0.4% ↓

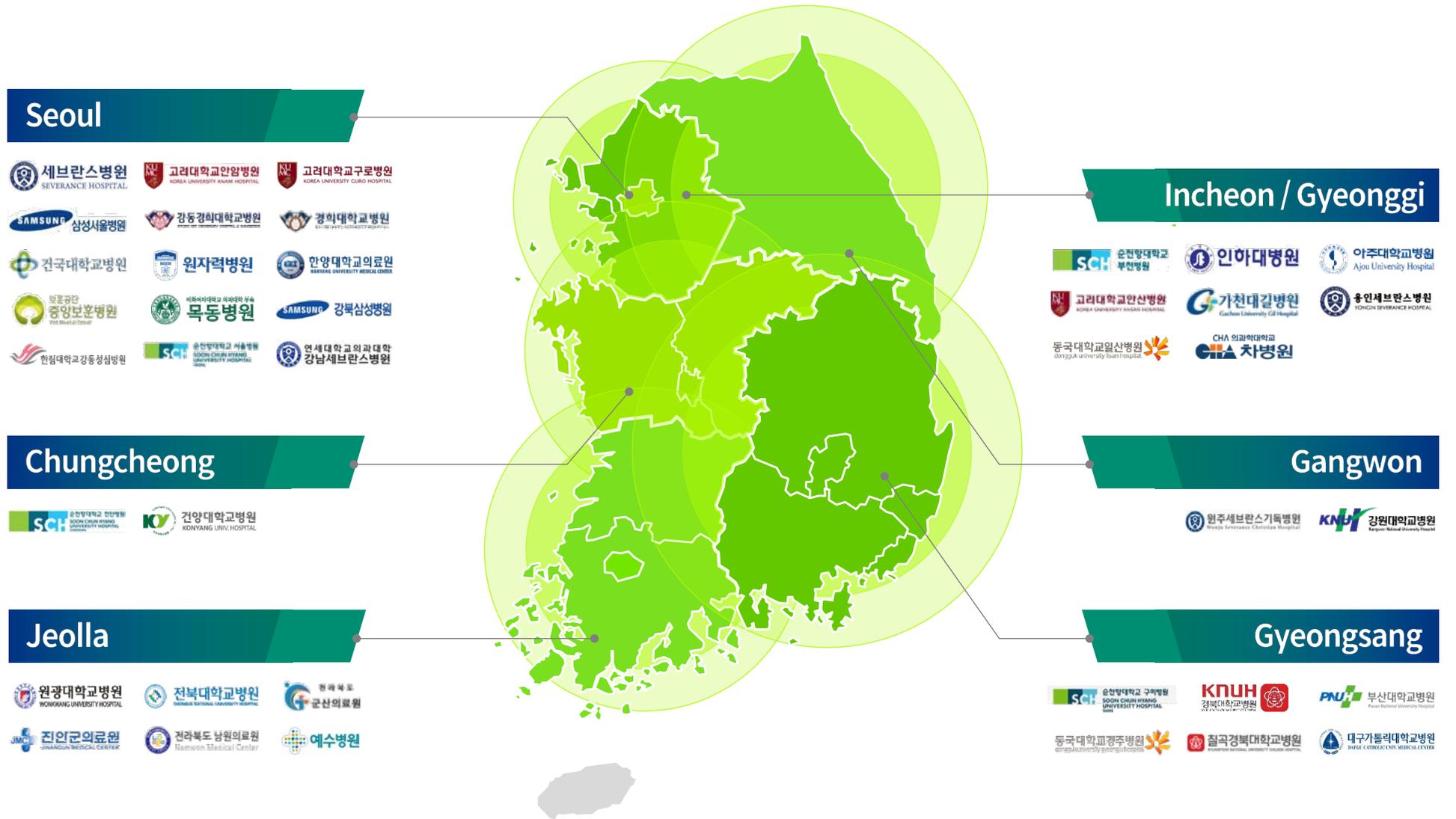# Korean Health Insurance System

Hospitals in Korea

- About 43 University or Tertiary Teaching Hospitals
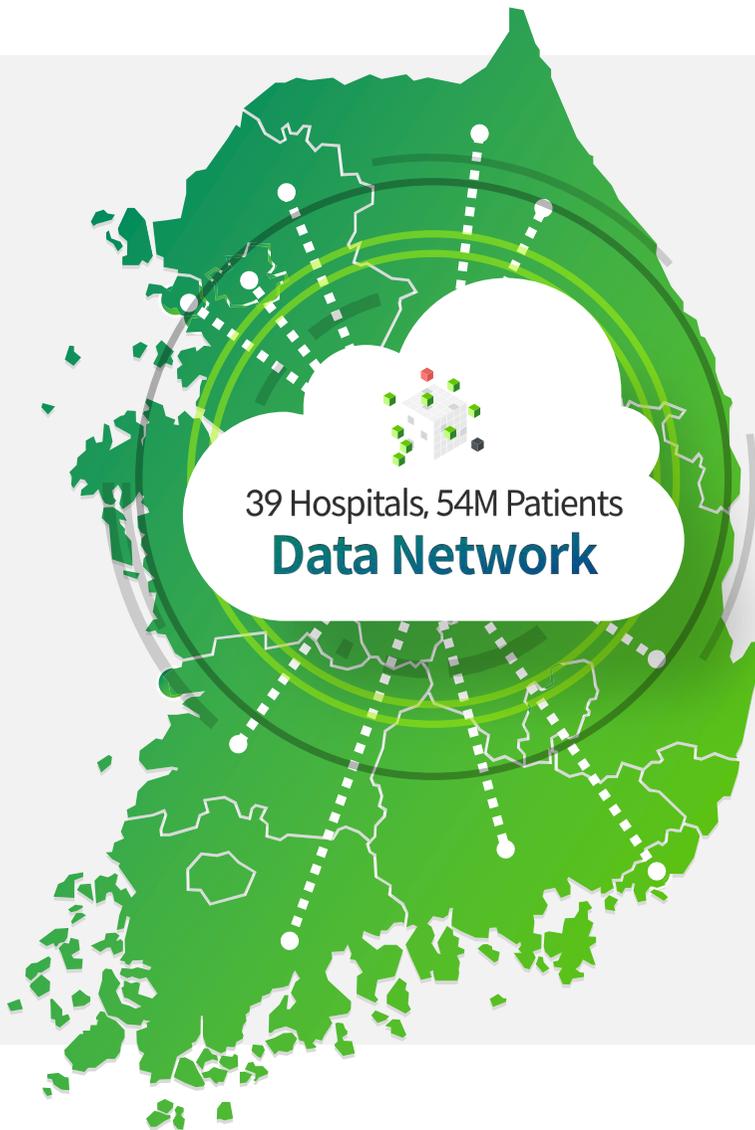- 280 general hospitals (>100 beds)
- 2,800 hospitals (>30 beds)

Status of
Korean OHDSI

Status of Korean OHDSI Network

AJOU UNIVERSITY | OHDSI

Data Network of 39 Hospitals, 54M Patients

Seoul

세브란스병원 SEVERANCE HOSPITAL | 고려대학교안암병원 KOREA UNIVERSITY ANAM HOSPITAL | 고려대학교구로병원 KOREA UNIVERSITY GURO HOSPITAL
SAMSUNG 삼성서울병원 | 강동경희대학교병원 | 경희대학교병원
건국대학교병원 | 원자력병원 | 한양대학교의료원 HANYANG UNIVERSITY MEDICAL CENTER
중앙보훈병원 | 목동병원 | SAMSUNG 강북삼성병원
한림대학교강동성심병원 | SCH 순천향대학교 서울병원 | 연세대학교 의과대학 강남세브란스병원

Incheon / Gyeonggi

SCH 순천향대학교 부천병원 | 인하대병원 | 아주대학교병원 Ajou University Hospital
고려대학교안산병원 KOREA UNIVERSITY ANSAN HOSPITAL | 가천대길병원 Gachon University Gil Hospital | 용인세브란스병원 YONGIN SEVERANCE HOSPITAL
동국대학교일산병원 dongguk university ilsan hospital | CHA 의과학대학교 차병원

Chungcheong

SCH 순천향대학교 천안병원 SOON CHUN HYANG UNIVERSITY HOSPITAL | 건양대학교병원 KONYANG UNIV. HOSPITAL

Gangwon

원주세브란스기독병원 Wonju Severance Christian Hospital | KNUH 강원대학교병원 Kangwon National University Hospital

Jeolla

원광대학교병원 WONKWANG UNIVERSITY HOSPITAL | 전북대학교병원 CHONBUK NATIONAL UNIVERSITY HOSPITAL | 전라북도 군산의료원
진안군의료원 JINANGUN MEDICAL CENTER | 전라북도 남원의료원 Namwon Medical Center | 예수병원

Gyeongsang

SCH 순천향대학교 구미병원 SOON CHUN HYANG UNIVERSITY HOSPITAL GUMI | KNUH 경북대학교병원 | PNU 부산대학교병원 Pusan National University Hospital
동국대학교경주병원 dongguk university gyeongju hospital | 칠곡경북대학교병원 KYUNGPOOK NATIONAL UNIVERSITY CHILGOK HOSPITAL | 대구가톨릭대학교병원 DAEGU CATHOLIC UNIV. MEDICAL CENTER

Commitment by 39 Hospital Heads to Share IRB Review Results and Conduct Joint Research

39 Hospitals, 54M Patients
Data Network

CDM Conversion History

2014
Ajou University Hospital
CDM Complete (First in Korea)

2015
Gachon University Gil Hospital
CDM Complete

2016.10
Korea National Health Insurance Service (NHIS)
CDM Conversion of 1M Patients

2017.09
Korea National Health Insurance Service (NHIS)
CDM Conversion of 1M Patients

2017.10
Gangwon State University Hospital
CDM Complete

2017.11
Korea Health Insurance Review and Assessment Service (HIRA)
CDM Conversion of 1M CV Patients

2018.01
Samsung Medical Center, Wonkwang University Hospital,
Wonju Severance Christian Hospital, Chonbuk National University Hospital
CDM Complete (4 Hospitals)

아주대학교병원 Ajou University Hospital

SAMSUNG 삼성서울병원

h·well 국민건강보험

가천대 길병원 Gachon University Gil Medical Center

KNU 강원대학교병원 Kangwon National University Hospital

연세대학교 원주기독병원 WONJU CHRISTIAN HOSPITAL

HIRA 건강보험심사평가원

원광대학교병원 WONKWANG UNIVERSITY HOSPITAL

전북대학교병원 CHONBUK NATIONAL UNIVERSITY HOSPITAL
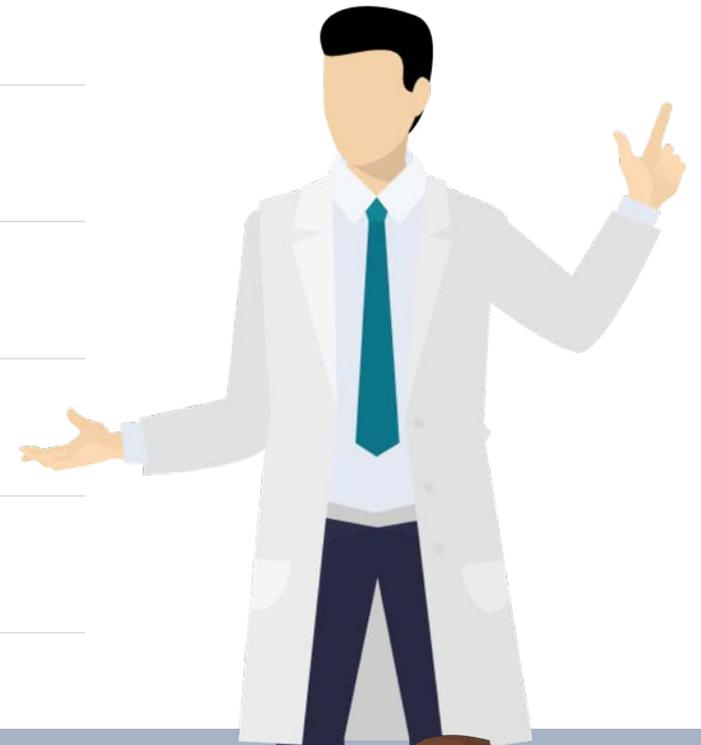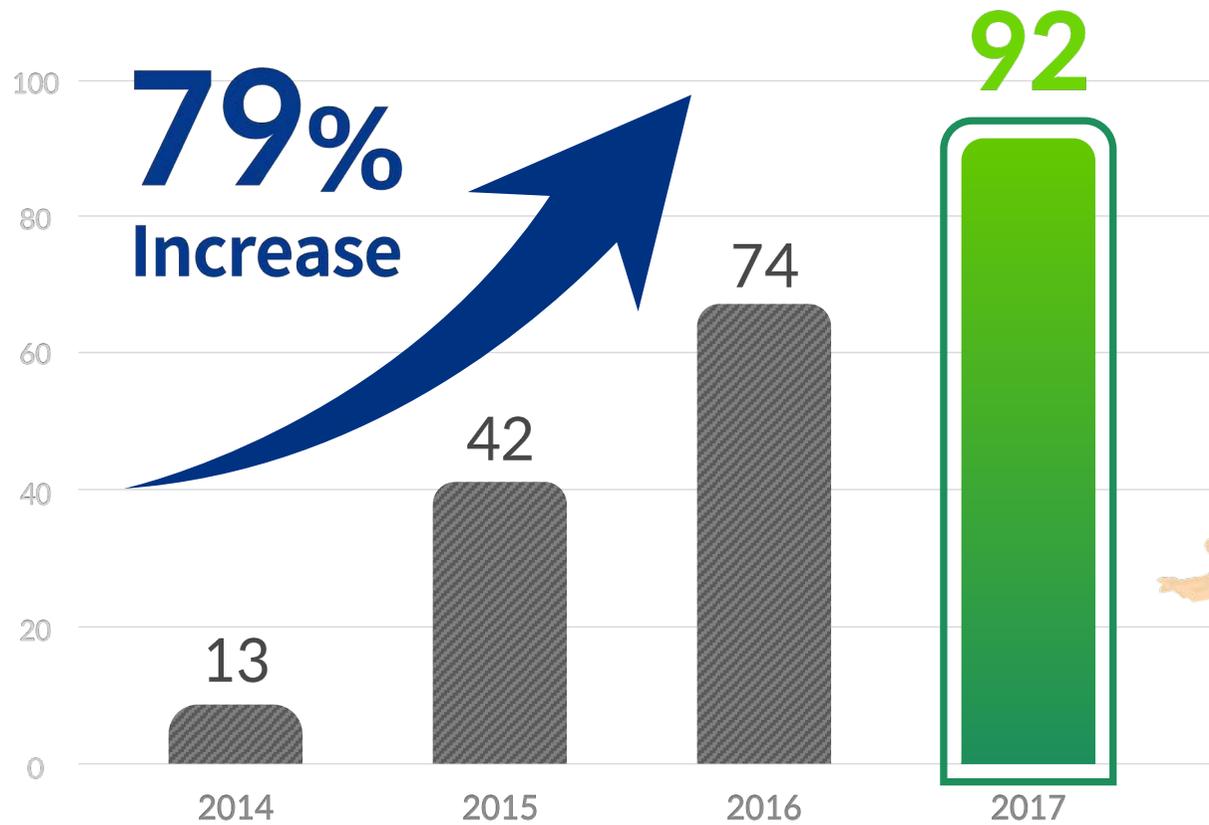
Why is **DRN** and **CDM** popular in Korea?

Invited Talks

226 invited talks during
past 42 months since July 2014

- 2014 : 13 times
- 2015 : 42 times
- 2016 : 74 times
- 2017 : 92 times

Number of invited presentations

No. of Presentations and Data Partner Join

● Presentations ▮ Join

Data Partners

Presentations

18 Months

200

12 Gacheon

76 NHIS

150

100

50

6
'14

11 12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12 1 2 3 4 5 6 7 8 9 10 11 12
'15   '16   '17

25

20

15

10

5

AJOU UNIVERSITY | OHDSI

Lesson learned from potential Data Owner

- Quick-prototyping
- Live demonstration
- Success story
- Focusing on clinicians
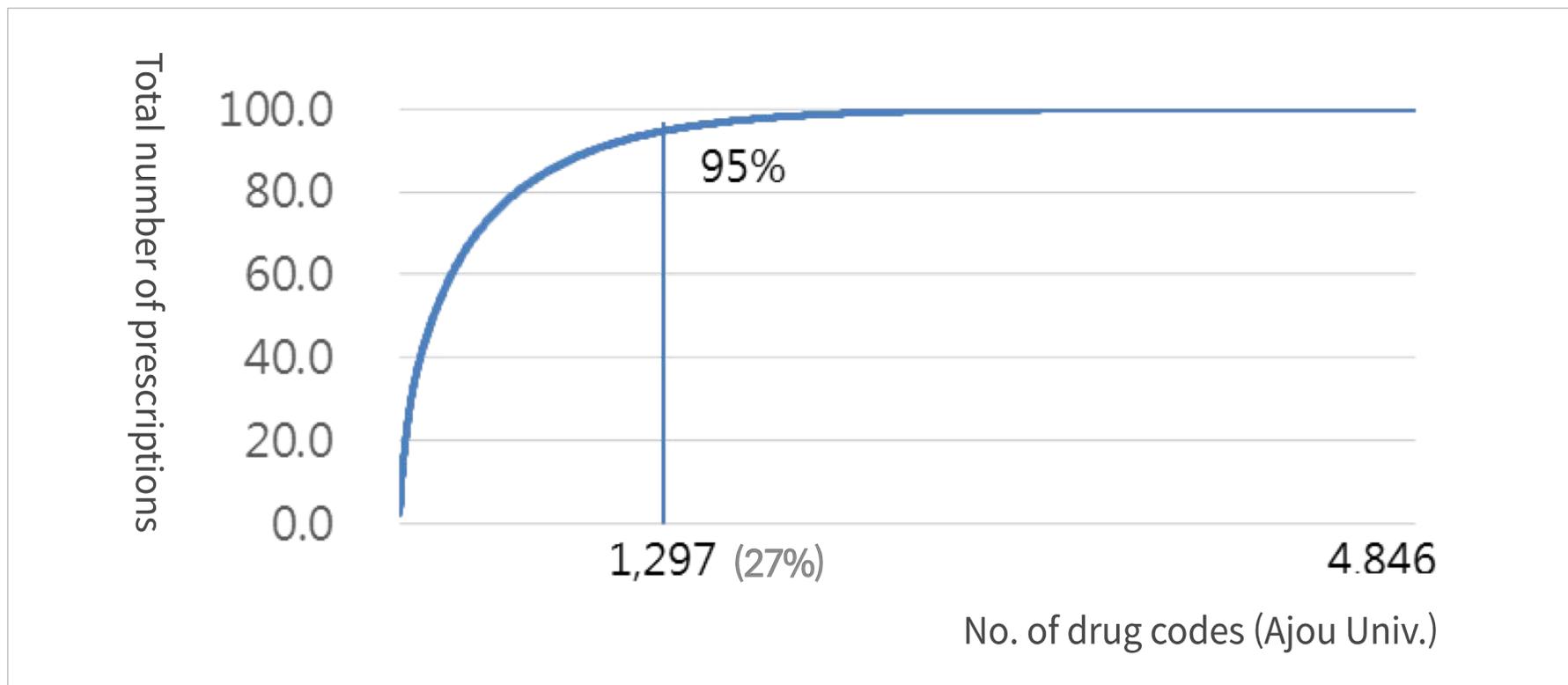- Data governance leadership meeting

Strategy for Standard Vocabulary Mapping

AJOU UNIVERSITY | OHDSI

- **Sort codes by frequency of usage**
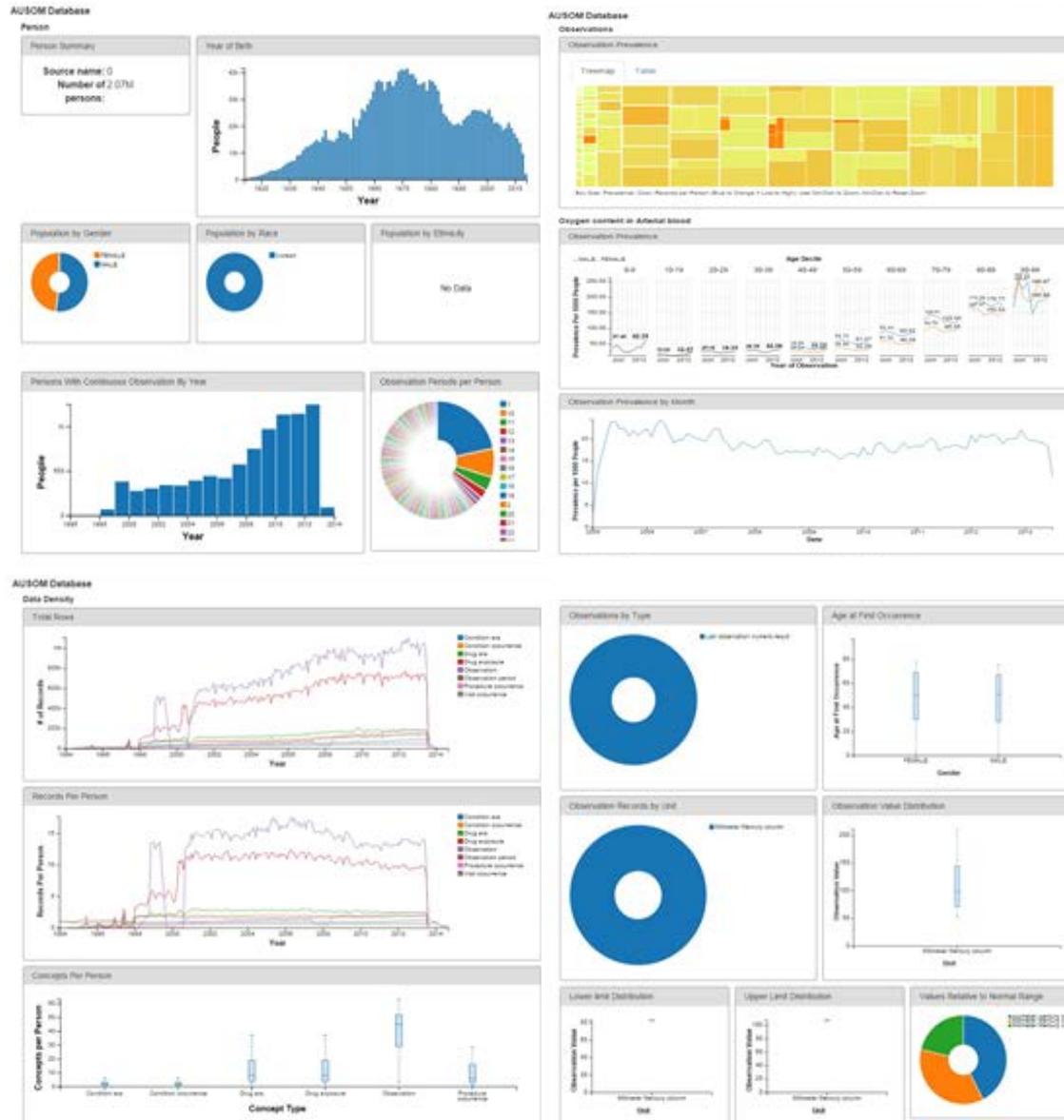- **Number of codes required to cover**

  95% of transaction data   ➡   **99%**
  20-30% of codes   **60-70%** of codes

Achilles : data characterization

Distributed research network

Openness
http://ami.ajou.ac.kr:8080

Examples of clinical researches using big data
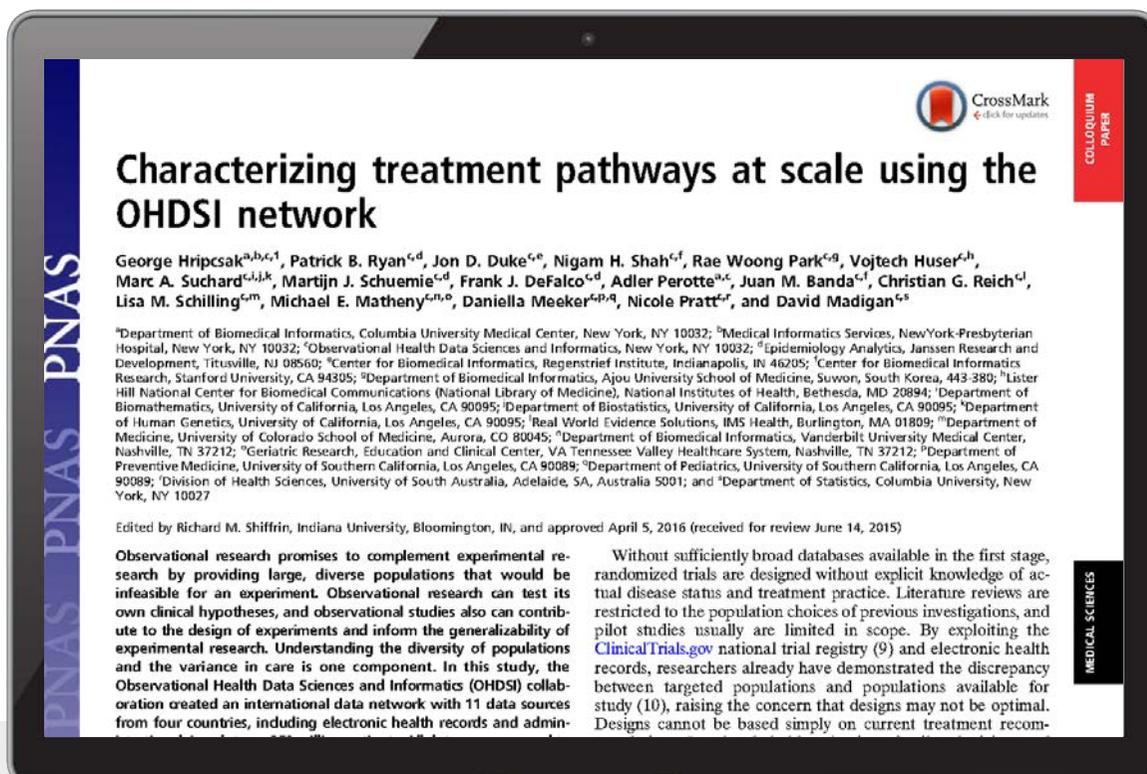
250M patients

12 database
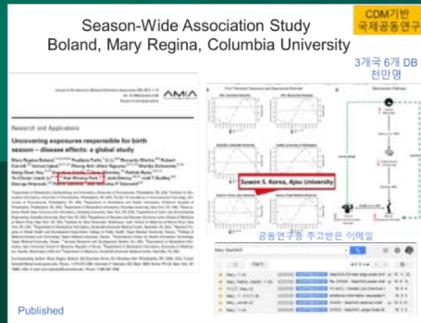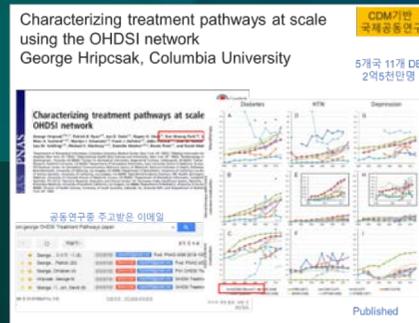
5 countries

Lead/participation in
CDM-based international joint researches (>14)

3 Countries, 6 DBs, 10M Patients

5 Countries, 11 DBs, 250M Patients

2 Countries, 24 DBs
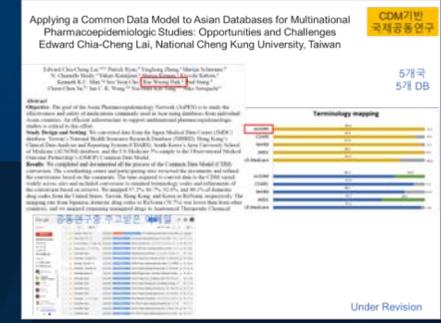
2 Countries, 24 DBs
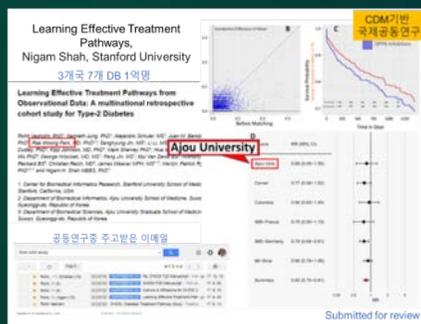
3 Countries, 7 DBs, 100M Patients
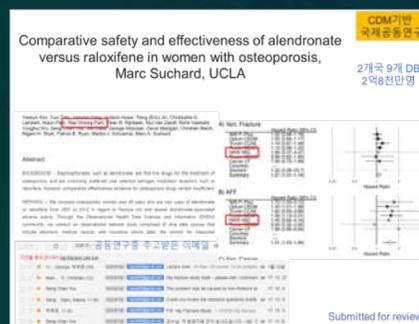
2 Countries, 9 DBs, 280M Patients
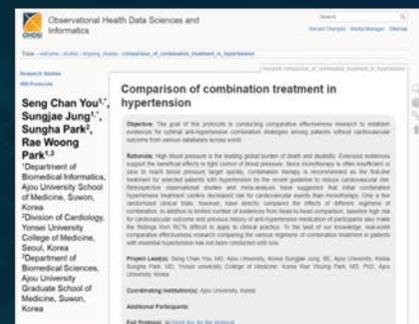
Hypertension Research Led by Ajou University

2017 OHDSI Symposium Best Practice Award

Focusing on clinicians

Clinicians, rather than informaticians

major decision makers
in a hospital are
usually clinicians

Young clinicians
assistant –
associate professors

highest interest on
OMOP CDM, because
they do not have enough
fund, resources and data
for their research

Inter-disciplinary clinical meeting

then homogenous
clinical meeting at
initial stage

Data Governance Leadership Meeting

AJOU UNIVERSITY | **OHDSI**

- Industrial Growth
- Information Security
- Promote Innovation
- Research Ethics
- Efficiency
- Privacy Protection

Ajou University has been coordinating OHDSI in Korea since March 2017 and is **actively initiating democratic healthcare data governance discussions** with CIOs from hospitals of various sizes

Ajou University Hospital
2017.03.07

Samsung Medical Center
2017.05.31

Gangdong Sacred Heart Hospital
(Hallym University) 2017.07.27

Ewha Womans University Mokdong Hospital 2017.09.21

Korea University Anam Hospital
2017.11.09

Korea University Guro Hospital
2018.01.25

Activity in Korea

Engineer Meeting

- Biweekly, TC (current 19th)
- EHR experts from participating hospitals
- Discuss all the technical issues during CDM conversion

Open Forum

- Monthly, 1-3 hour lecture
- Agenda
  - Introduction to OHDSI and CDM
  - OMOP CDM Structure
  - OMOP CDM Vocabulary/ vocabulary mapping
  - Tools for OMOP CDM
  - ETL process
  - Research Experience using OHDSI Network

Government's interest in CDM

No. of Presentations and Data Partner Join

NHIS, Development of a CDM-based Drug Safety Surveillance System

AJOU UNIVERSITY | OHDSI

NHIS

Claim data
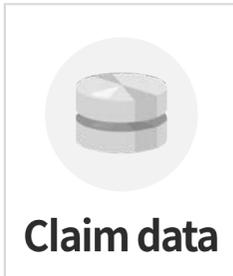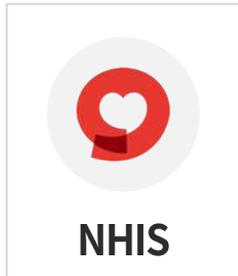
Three year project 2016-2018

1st year | **Feasibility**
9-year of 1M pt data into CDM

2st year | **Validation**
validation of usefulness of the CDM

3st year | **Full conversion**
12-years of 51 M patients

OMOP CDM conversion

Application of pharmacovigilance tools

RAW data
SAS file?

ETL Layer
Voca mapper
ETL
Data Conversion

Standard Vocabularies
SNOMED-CT
RxNorm
LOINC

DQM
Analyses
Evaluation
Desing
Conversion

PV tools
ROR, PRR, LGPS
CLEAR CERT

OHDSI tools
ATLAS
ACHILLES
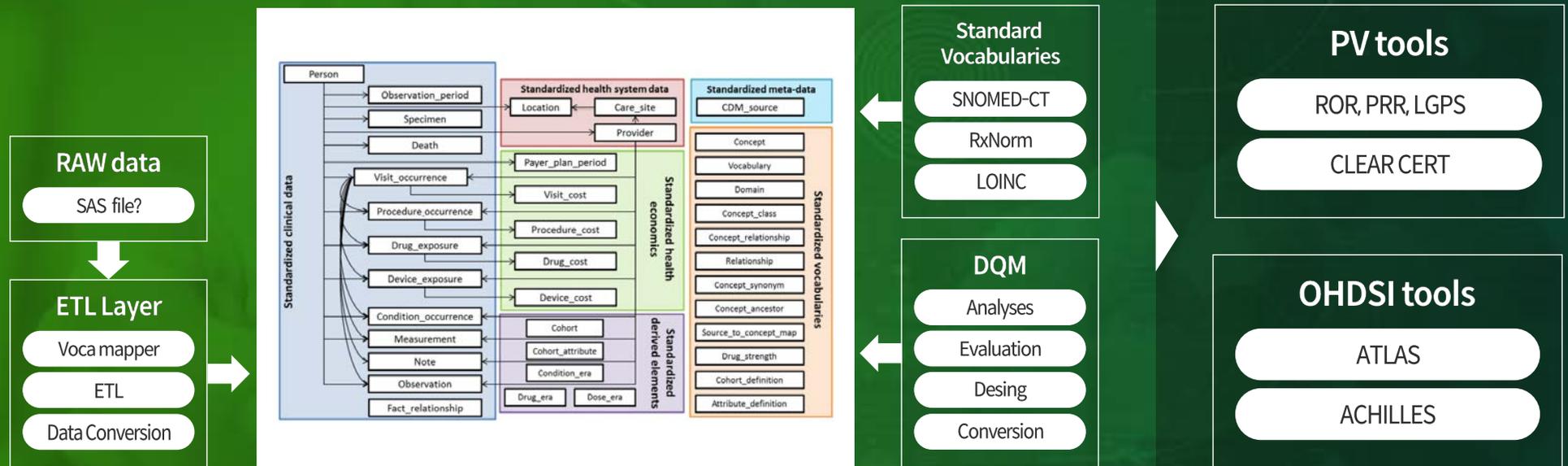
National Health Insurance Service : **Governmental National Health Insurer**

Project Results

Designing an Efficient System to Support CDM Conversion
of the Entire Korean Population

Simulation of required system specifications
(based on Oracle Exadata)

- Uncompressed CDM data of 1M patients (2%) = 527GB
- Uncompressed CDM data of 50M patients (100%) = 25.8TB (estimate)
- Total 103TB (x4) required considering index and temp domain

Table Name	Column Size (Byte)	Data Count	Total Size (MB)
Patient Information	261	999,812	248.9
Visit Information	152	224,085,191	32,483.10
Location Information	400	317	0.1
Site Information	406	188,814	73.1
Drug Exposure Information	344	406,670,758	133,414.00
Continuous Drug Exposure Period	10	154,502,958	1,473.50
Diagnosis Information	181	355,902,525	61,434.10
Continuous Diagnosis Period	10	125,949,247	1,201.10
Observation Information	521	186,289,956	92,560.80
Death Information	506	45,767	22.1
Observation Period Information	9	999,813	8.6
Insurance Coverage Period	158	999,812	150.7
Procedure Information	160	961,582,127	146,725.80
Measurement Information	283	260,294,988	70,251.00
Total	3,401	2,678,512,085	540,046.80

Project Results

Drug Use Analysis System Based on NHIS CDM Data

⤬ Roadmap

	CDM Infrastructure	CDM Analysis System
Level 1 (2016)	**Pilot** • Development of claim data CDM conversion technology • 1M sample data conversion	**Pilot** • Causal inference of drug-induced AEs
Level 2 (2017)	**Validation** • Code mapping update and data quality management • Calculation of required system specifications	**Validation** • Validation of analysis system and tool feasibility • Improvement of causal inference analysis
Level 3 (2018~)	**Scale-up** • CDM conversion of entire Korean population (50M) • CDM table extension • System environment setup	**Launch** • Automation of large-scale analysis • Trend analysis of data usage • Monitoring of disease indicators • Cost effectiveness analysis

Complete

Level 4 (2018~) — **Expansion of CDM Data Application**

- Promotion and training
- Big data analysis environment setup
- Development and application of various research methodologies
- Development of tools such as drug-induced AE prediction, visualization of analysis results, etc.
- Continuous CDM maintenance

Drug Use Analysis System Based on NHIS CDM Data

CDM Applications

Information for Public

External
Research Assistance

h·well
NHIS National Health Insurance Service

Application within NHIS

- Provide resources related to drug pricing
- Generate background data to improve reimbursement
- Big data analysis training

External Research Assistance

- Research on therapeutic effect vs. AE
- Usage analysis of drugs, etc.
- Personalized disease predictions
- Reference data for policy research
- Reference data for new drug development
- Research on international disease and treatment trends
- Development of research methodologies and analytic tools

Information for Public

- Provide information about drugs, disease, etc.
- Provide information on potential high risk groups
- Provide information on reimbursed drugs

The Korea Herald

THE
INVESTOR
March 12, 2018

Bio & Medicine
Korea to build database for medical big data

PUBLISHED :February 09, 2018 - 17:13   UPDATED :February 09, 2018 - 17:13

A A f y ✉ 🖶

[THE INVESTOR] The South Korean government will consolidate medical data from hospitals across the country and establish a comprehensive database, as part of its goal to foster the country's lucrative pharmaceutical and healthcare industries, officials said on Feb. 9.

Its extended vision is to create some 35,000 new jobs in the growing bio and healthcare sector and to pull up the nation's global market share from the current 1.8 percent to 4 percent by 2022.

The Ministry of Trade, Industry and Energy on Feb. 9 held a conference with local bio & healthcare players at Wonju Medical Industry Techno Valley and announced the government's short- and mid-term development strategies for the industry.

The event was attended by officials from Daewoong Pharm, Mediana Samsung Seoul Hospital, Korea Biotechnology Industry Organization and other local medical organizations.

As a pilot step, the ministry is to select six major hospitals within the first half of the year and collect genetic and biometric data of some 10 million people to create a bio database on bio information, according to officials.

Distributed Bio Big Data Model

Bio Big Data Center

- Analysis SW ← Hospital
- Analysis Results → Hospital
- Analysis request ← Company
- Analysis Results → Company
- Analysis SW request ← IT Company
- Analysis SW supply → IT Company
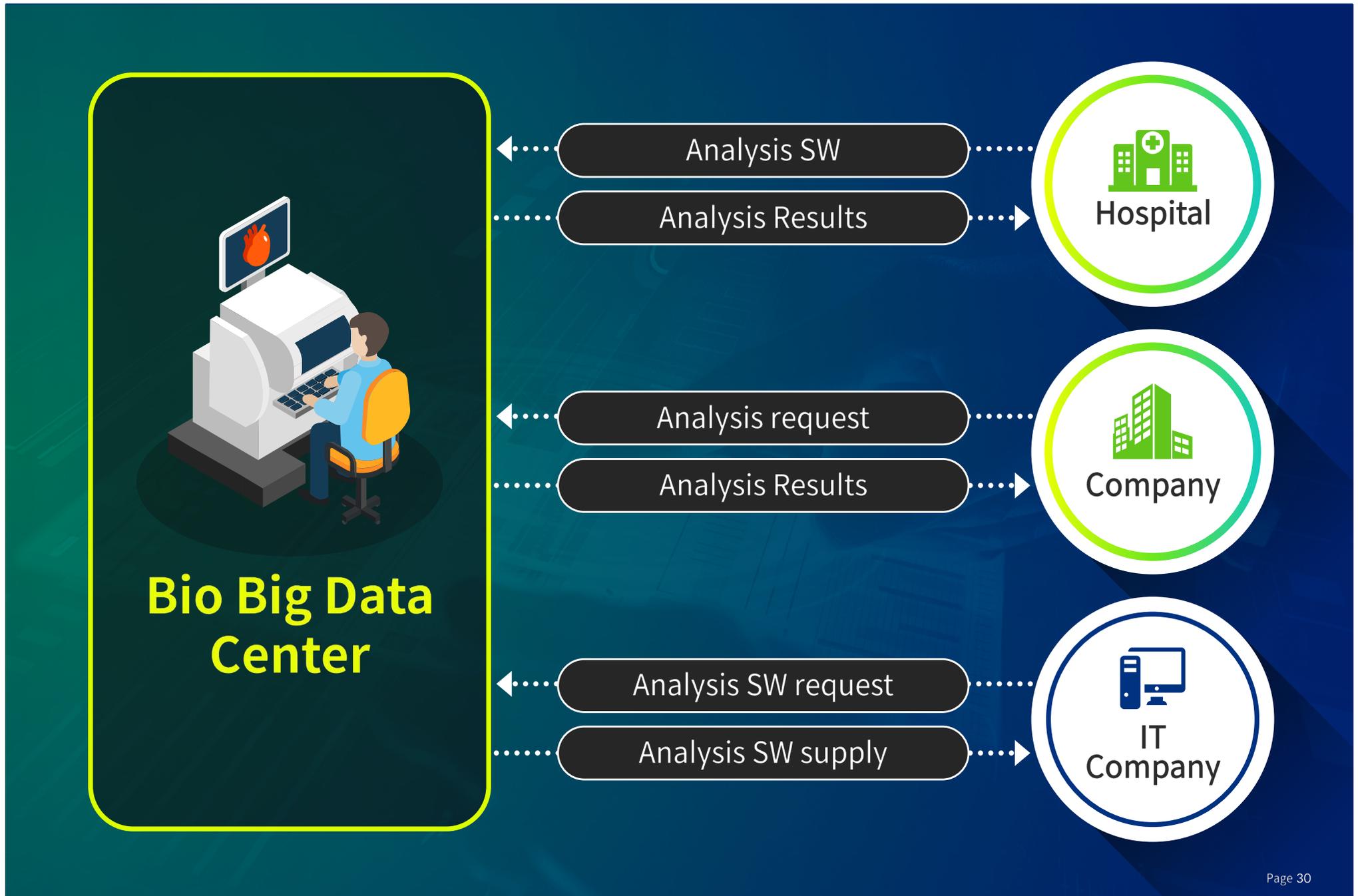
News > Economy > General

Ministry of Industry - Ministry of Health and Welfare 'Bio-Big Data'
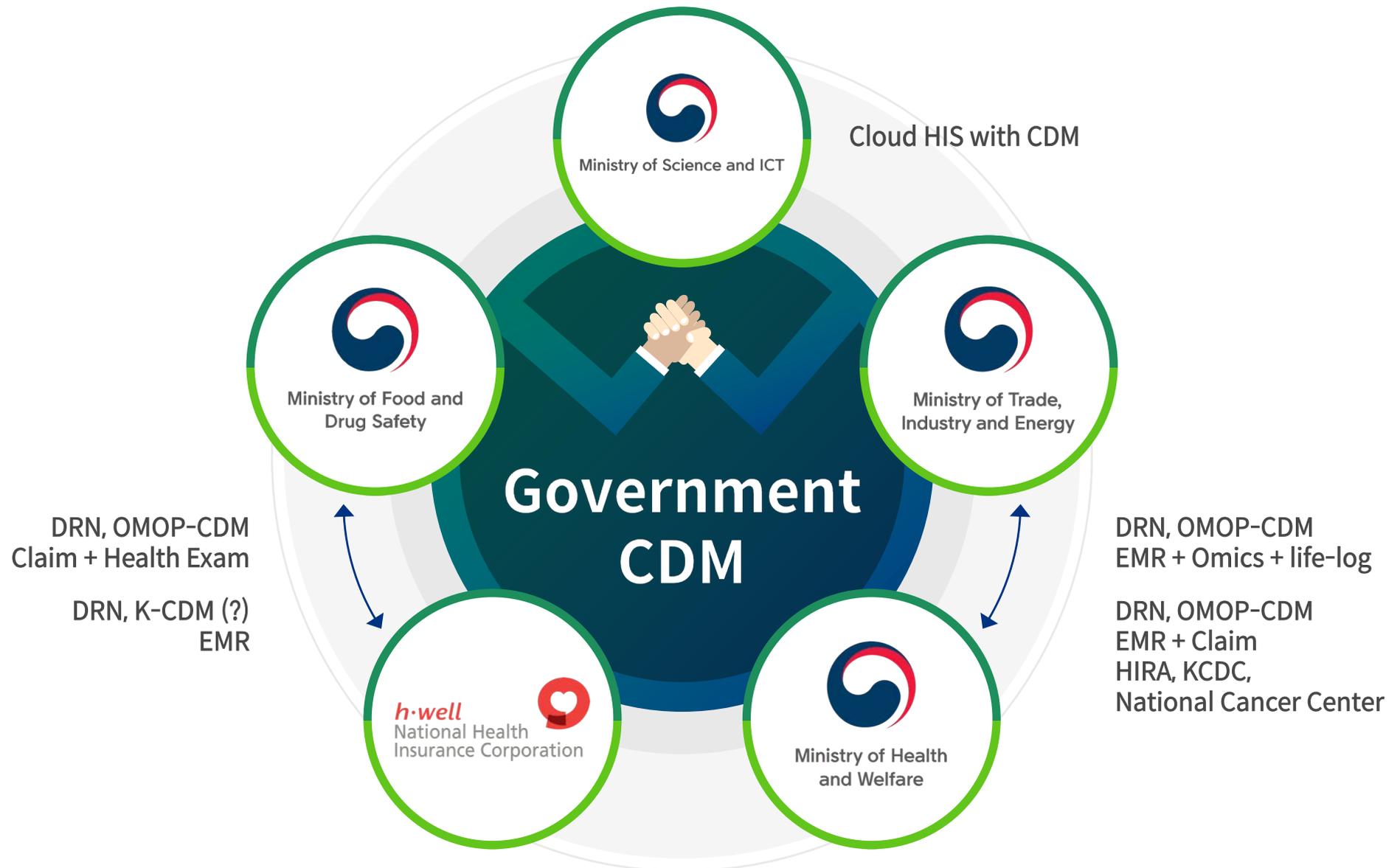
Enter the article: 2017-04-20 10:16

[Election Day Sejong = Park Eally]

Ministry of Health and Welfare launches big medical data promotion group in March ... The Ministry of Industry is also promoting TF

In the era of the Fourth Industrial Revolution, in order to preoccupy the bio-data field, there is a fierce nervous battle between government departments.

According to the Ministry of Health and Welfare, the Ministry of Health and Welfare, which is in charge of the hospitals and the pharmaceutical industry, started a pilot project last year to build and utilize healthcare data. Also, last month, the 'Big Health Data Medical Promotion Team' was launched in order to create a strategy to utilize big medical data. The Ministry of Industry, Trade and Industry and the Future Creation Science Department also participated in the promotion team, and the head of the Health and Medical Policy Division of the Kim Gang - The Ministry of Health and Welfare is planning to establish a platform to link big data of public institutions such as National Health Insurance Corporation and Health Insurance Review and Evaluation Center.

Collaboration and Competition in between Government CDM

Cloud HIS with CDM

Ministry of Science and ICT

Ministry of Food and Drug Safety

Ministry of Trade, Industry and Energy

Government CDM

DRN, OMOP-CDM
Claim + Health Exam

DRN, K-CDM (?)
EMR

h·well
National Health
Insurance Corporation

Ministry of Health
and Welfare

DRN, OMOP-CDM
EMR + Omics + life-log

DRN, OMOP-CDM
EMR + Claim
HIRA, KCDC,
National Cancer Center

Characteristics of Korean OHDSI

Data partners

- Major tertiary teaching hospitals
  - Detailed time stamp
  - Test results
  - Outcome data

+

National Health Insurance Data

- Compulsory health insurance

- Claim data
  + socioeconomic data
  + regular Health exam data (includes lab tests)

- 10-year of observation period

- Covers all the citizens (50M)

KOHDSI
KOREA OBSERVATIONAL HEALTH DATA SCIENCES AND INFORMATICS

Plan in KOREA

CDM-Based

Distributed Data Network

Development of an OMOP CDM-Based Distributed, Integrative Bio Health Data Network

AJOU UNIVERSITY | OHDSI
Observational Health Data Sciences and Informatics

Request for Proposal Overview

AJOU UNIVERSITY | OHDSI

Final Goal

- To develop key technologies required to setup and operate an OMOP CDM-based data network
- To promote industrial application of hospital data such as clinical data, genomics data, biosignal data, etc.

- **Development and validation of CDM extensions for genomics and biosignal data**
- **Conversion of EMR data to OMOP CDM**
- **Quality management of CDM data**
- **Development of analytical tools for CDM data**
- **Development of information security technologies**

Period **Within 36 months** (9 months in 2018)

Grant **Total $4M** ($0.8M in 2018)

Host Organization **No restrictions**

* Participation of hospitals and small/midsize businesses required

❖ Participating hospitals must agree to share IRB review results and conduct joint research (commitment from hospital heads required)

❖ Participating businesses must present respective business models based on CDM

Development Goals

	Evaluation Criteria	Goal
1	Development of CDM extensions (genomics, lifelog, etc.)	≥2
2	Participating hospitals	≥6
3	Patients converted to CDM	≥1M
4	Joint researches	≥5
5	Business models based on CDM data	≥3
6	Development of data quality assurance tools for CDM	≥3

Proposal Outline

Hospitals
(39, 34K beds)

Patient Data (54M)

Extension Model (7)

Result

Query

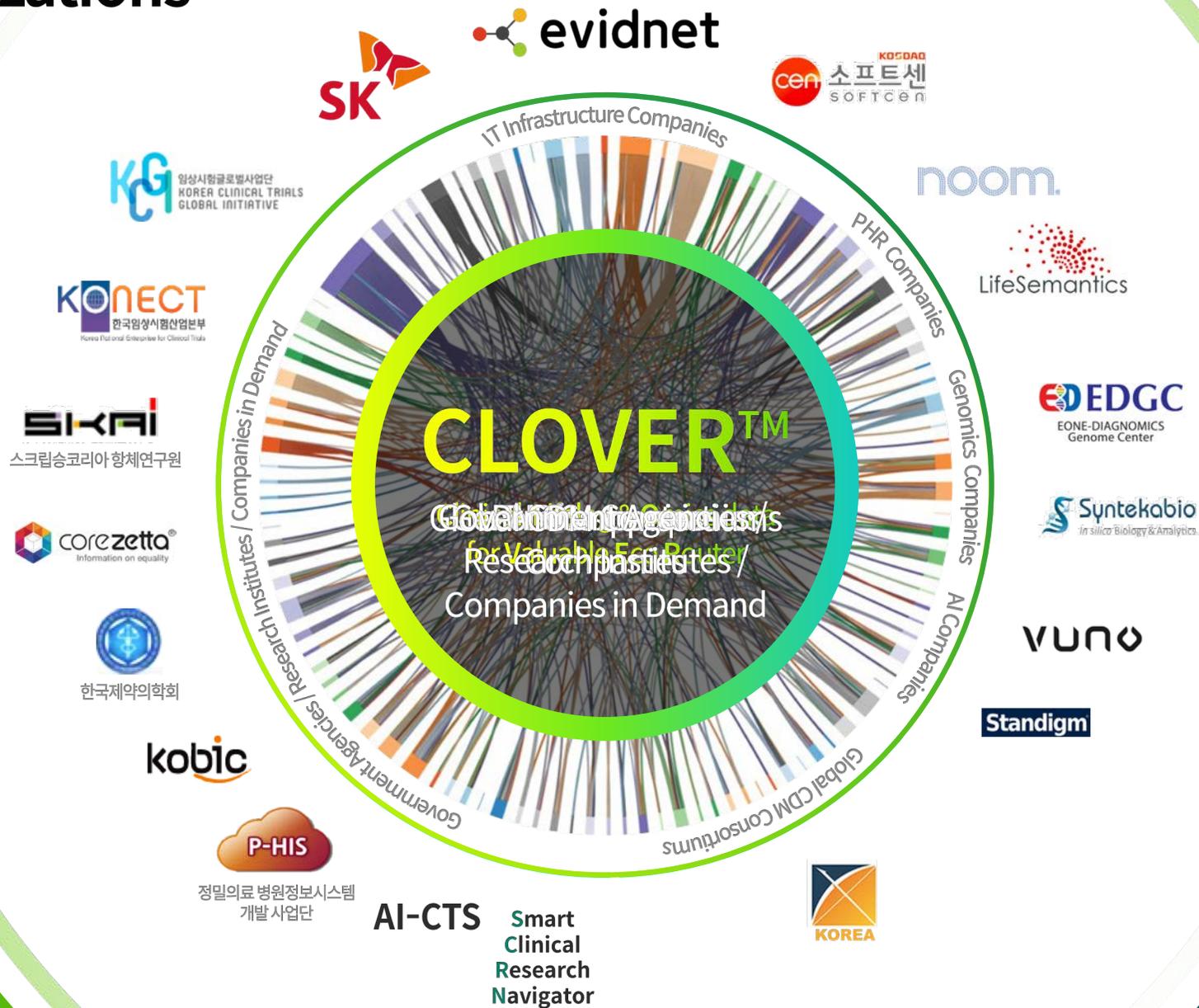## Coordinating Platform

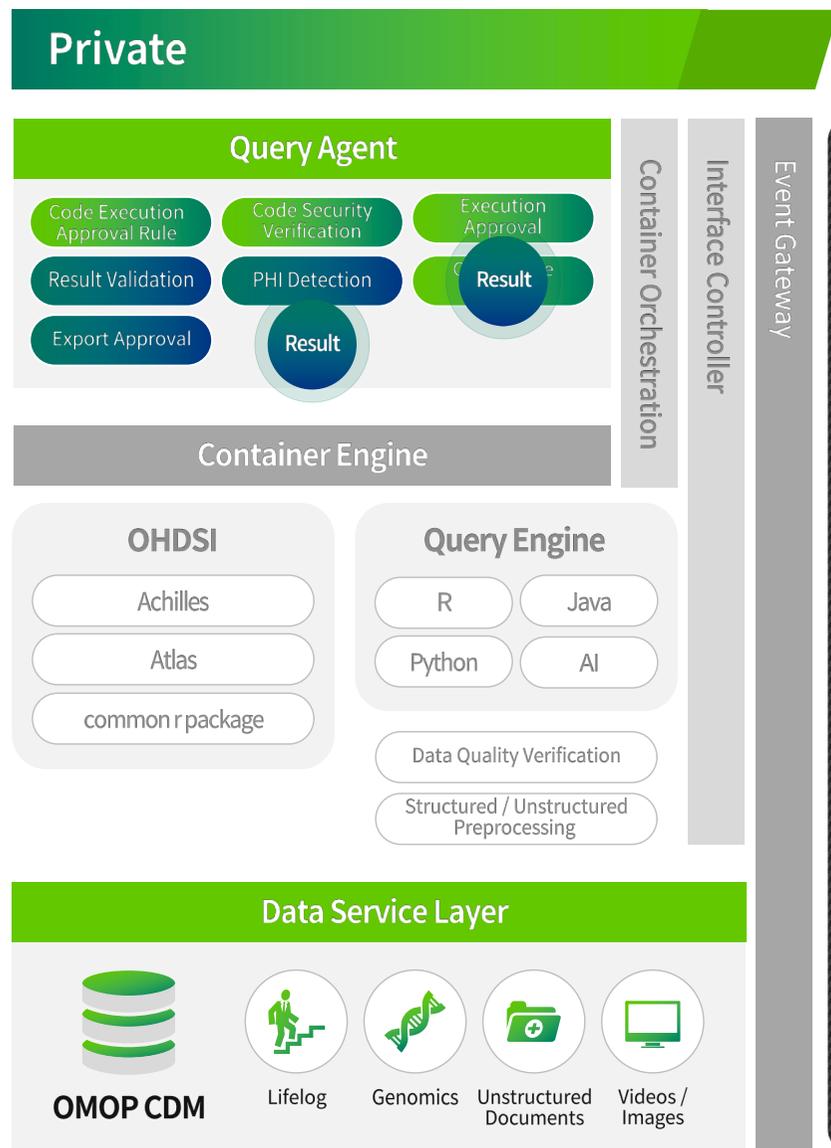**Research Institutes** (10)

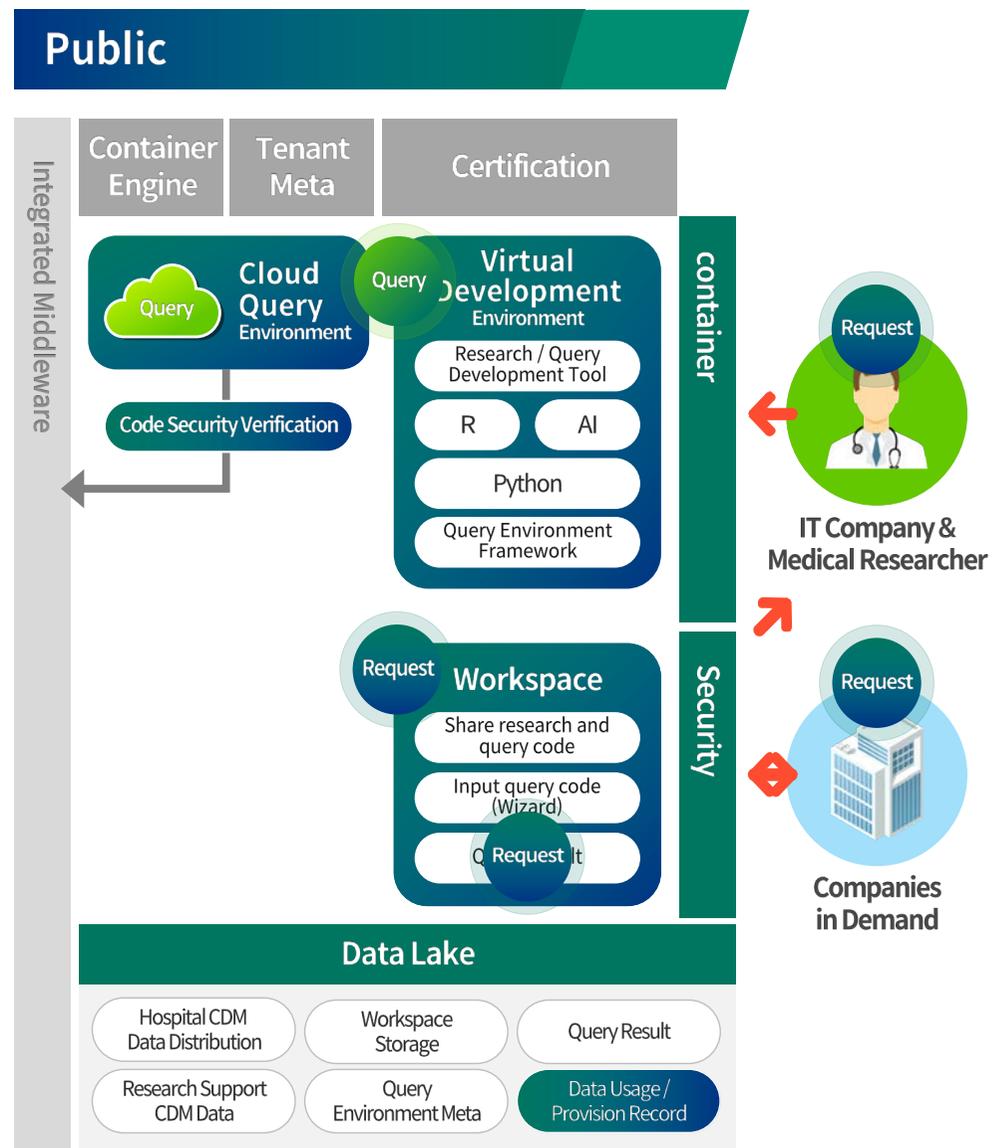**Companies** (7)

BM (7)

Collaborating or Participating Organizations

CLOVER™ Platform

Clinical, Lifelog & Omics Data for Valuable Eco Router

CLOVER™ Node (Participating Hospital)

Private

Query Agent
- Code Execution Approval Rule
- Code Security Verification
- Execution Approval
- Result Validation
- PHI Detection
- Result
- Export Approval
- Result

Container Orchestration

Interface Controller

Event Gateway

Container Engine

OHDSI
- Achilles
- Atlas
- common r package

Query Engine
- R
- Java
- Python
- AI
- Data Quality Verification
- Structured / Unstructured Preprocessing

Data Service Layer
- OMOP CDM
- Lifelog
- Genomics
- Unstructured Documents
- Videos / Images

Fire Wall

CLOVER™ Portal

Public

Integrated Middleware

- Container Engine
- Tenant Meta
- Certification

Cloud Query Environment
- Query

Virtual Development Environment
- Query
- Research / Query Development Tool
- R
- AI
- Python
- Query Environment Framework

Code Security Verification

Workspace
- Request
- Share research and query code
- Input query code (Wizard)
- Query Result / Request

container

Security

Data Lake
- Hospital CDM Data Distribution
- Workspace Storage
- Query Result
- Research Support CDM Data
- Query Environment Meta
- Data Usage / Provision Record

Request — IT Company & Medical Researcher

Request — Companies in Demand

PHI Protection

CLOVER™ Node (Participating Hospital)

Private

Query Agent
- Code Execution Approval Rule
- Code Security Verification
- Execution Approval
- Result Validation
- PHI Detection
- Query Code Execution
- Export Approval

Container Orchestration

Interface Controller

Event Gateway

Container Engine

OHDSI
- Achilles
- Atlas
- common r package

Query Engine
- R
- Java
- Python
- AI

Data Quality Verification

Structured / Unstructured Preprocessing

PHI detection

OMOP CDM Lifelog Genomics Unstructured Documents Videos / Images

Fire Wall

CLOVER™ Porta

Integrated Middleware

Container Engine

Tenant Meta

Cloud Query Environment

Query Checkup

Da...

Hospital CDM Data Distribution

Research Support CDM Data

We... S...

Enviro...

evidnet

01
Verification of Sensitive Data

Detection and de-identification of PHI in compliance to HIPAA

02
Verification of Code Security

Detection of malware prior to query code import

03
Detection of PHI

Detection of PHI based on k-anonymity, l-diversity and t-closeness prior to query result export

Personalized Services

AJOU UNIVERSITY | OHDSI

Hospital
EMR

OMOP / CDM

Company
Life-log   Genomic ...

DATA BASE   Data Service Layer

Data Interface
Platform

Deep Learning Engine

Analysis Engine

EDGC EONE-DIAGNOMICS Genome Center
Genomics CDM Extension Based
BM Discovery

Syntekabio An Asian Biology & Analytics
Genomic Biomarker and Resistance Mutation Based
Personalized Oncology
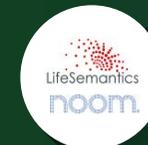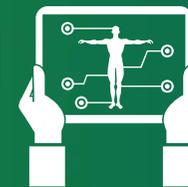
LifeSemantics noom
Chronic Disease High-Risk Group
Care Service

Services based on genomic + clinical data

- Patent application of platform technology and service : >2
- Commercialization : >2

Chronic Disease Care Service

Perinatal Mother Care Service

Blood

Tissue
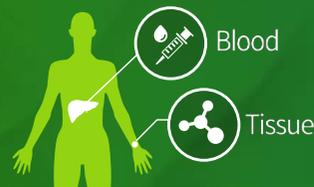
NGS Test and Data Analysis

Discovery of **Cancer-specific Mutations**

CDS Gallery
- Gene mutation-protein structure-drug correlation visualization
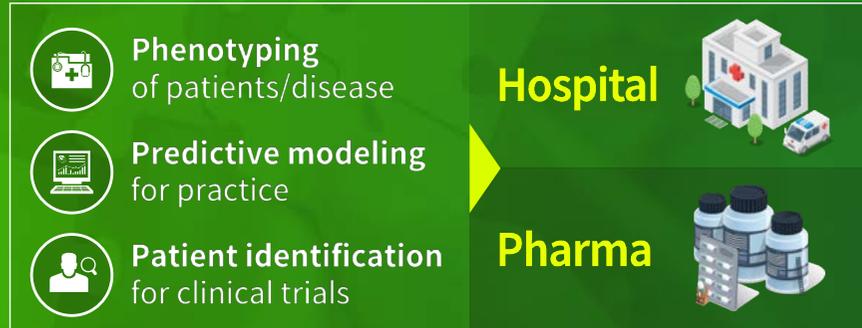- Treatment recommendation based on mutations

PHR Platform

OPEN API

Bio CDM Platform

진료, 검사, 영상기록 EMR Data

유전체 데이터 Gene Data

라이프로그 데이터 Lifelog Data

Dietary and prescription service
for chronic disease patients after discharge

Pharma R&D Services

Hospital
EMR

Company
Life-log **Genomic** ···

OMOP / CDM ← **DATA BASE** (Data Service Layer)

Data Interface
Platform

Deep Learning Engine

Analysis Engine

Standigm

Clinical Trial Eligible
Patient Identification

Deep EMR System

- Standigm DB
- AI Drug Development Model

- **Phenotyping** of patients/disease
- **Predictive modeling** for practice
- **Patient identification** for clinical trials

Hospital

Pharma

---

corezetta

Fixed Dose Combination (FDC) Recommendation &
Drug Repositioning

Corezetta
corezetta
Information on equality

CDM. Pharma System
Big data analysis, machine learning

Pharma

| Drug-induced AE analysis | Co-prescription pattern and patient outcome analysis | Patient outcome analysis based on prescription schedule | Discovery of unmet needs |

Results

- Incrementally modified drugs (IMDs) with improved safety profile
- New FDCs
- Discovery of niche markets
- Protocol optimization and feasibility assessment

Summary

01 **About South Korea**

02 **Korea OHDSI Network**

03 **Why is DRN and CDM popular in Korea?**

04 **Lesson learned from potential Data Owner**
- Quick-prototyping
- Live demonstration
- Success story
- Focusing on clinicians

05 **Future plans in Korean OHDSI**

Q&A

Contacts

Rae Woong Park

- E-mail : veritas@ajou.ac.kr
- Tel : +82-31-219-4471