# Digital Tools and Infrastructure:
## The Role of Artificial Intelligence in Open Science

Chee Sun LIEW

cheesunl@sunway.edu.my





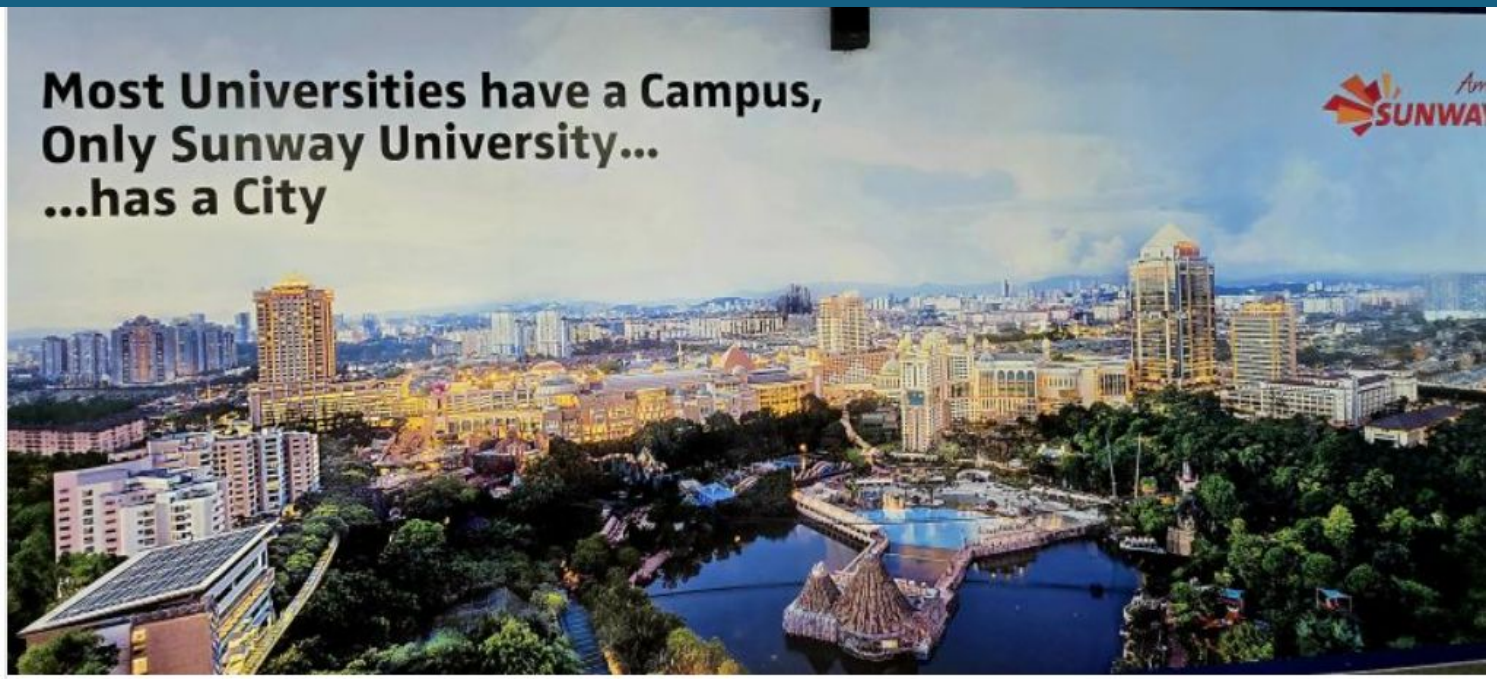

WASTELAND TO WONDERLAND
SUNWAY CITY
Lee Kuan Yew







# Sustainability is in the DNA of Sunway and Sunway University






# POWERED BY THE SUNWAY® ECOSYSTEM



Supported by the vast Sunway Group ecosystem, from healthcare and education to property and innovation. Sunway University thrives within a network designed to nurture holistic, real-world impact.

A complete ecosystem where ideas move seamlessly from lab to life.



PROPERTY



CONSTRUCTION



MALLS



TRADING & MANUFACTURING



QUARRY



DIGITAL



HOSPITALITY



LEISURE



HEALTHCARE



BUILDING MATERIALS




INVESTMENT & FINANCIAL SERVICES




REIT



EDUCATION




SUNWAY UNIVERSITY

# Sunway University Quick Facts

**14,280**

***STUDENTS***

**24% International**

**140**

***ACADEMIC PROGRAMMES***

**8**

***ACADEMIC FACULTIES/SCHOOLS***

**17**

***RESEARCH CENTRES***

**NOT FOR PROFIT**

***PRIVATE UNIVERSITY***

Sunway Business School, Faculty of Arts and Social Sciences, Faculty of Engineering & Technology, Faculty of Medical & Life Sciences, School of Hospitality & Tourism Management, School of Mathematical Sciences, School of American Education, School of Education



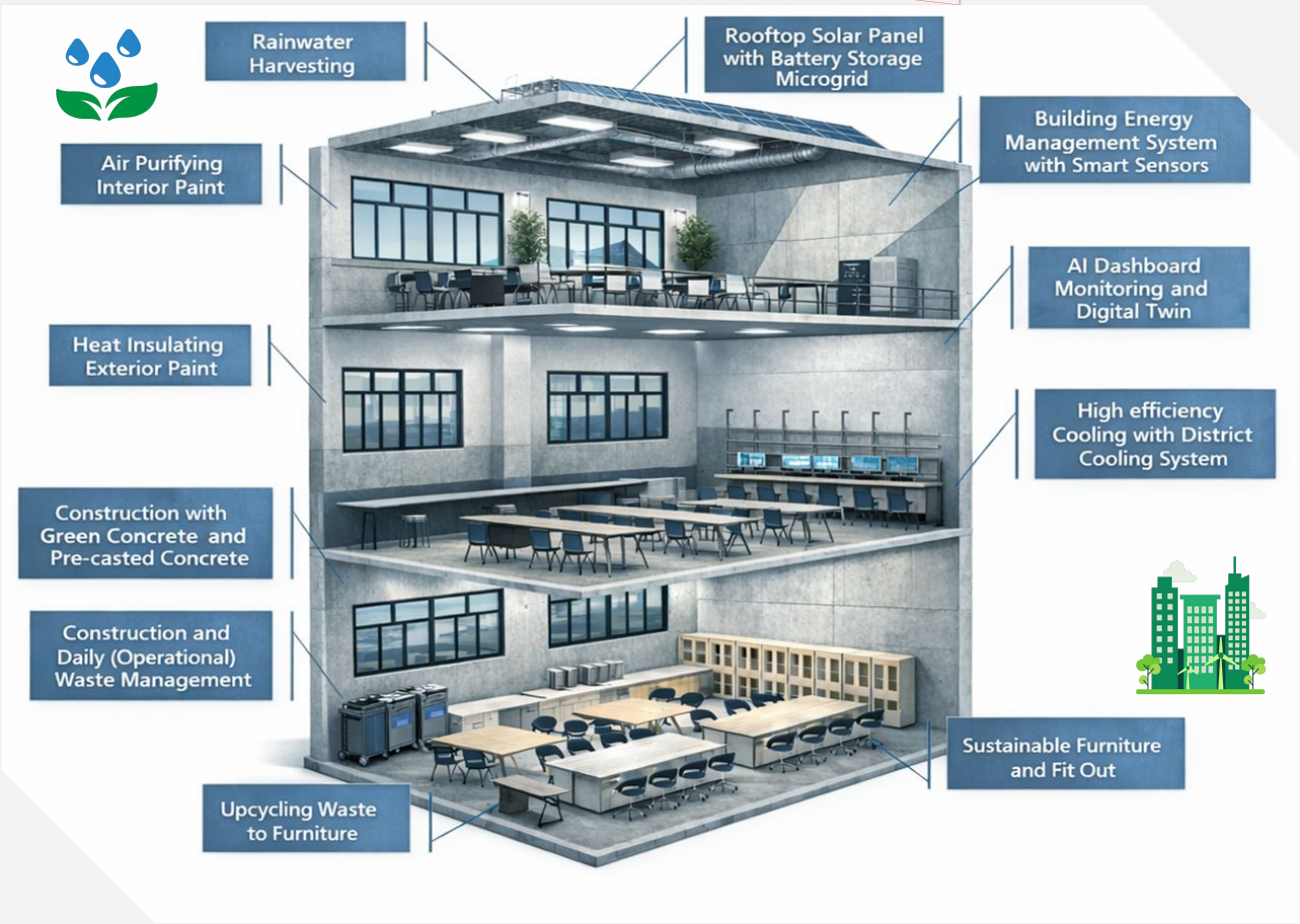
# A LIVING LABORATORY OF SUSTAINABILITY

Sunway City stands today as Malaysia’s living laboratory - a testbed for green technologies, smart systems, and sustainable lifestyles.

Here, the city is not just where we live. It is how we learn.

# NET ZERO CARBON BUILDING






What if the most valuable output of a research project is not just its publication, but the data, software and workflows that enable thousands of other discoveries?

# Let’s focus on these THREE questions

I’ll love to discuss more, but we only have 30 minutes

1. What is the digital foundation of modern research?
2. How does Open Science empower AI?
3. How does AI advance Open Science?

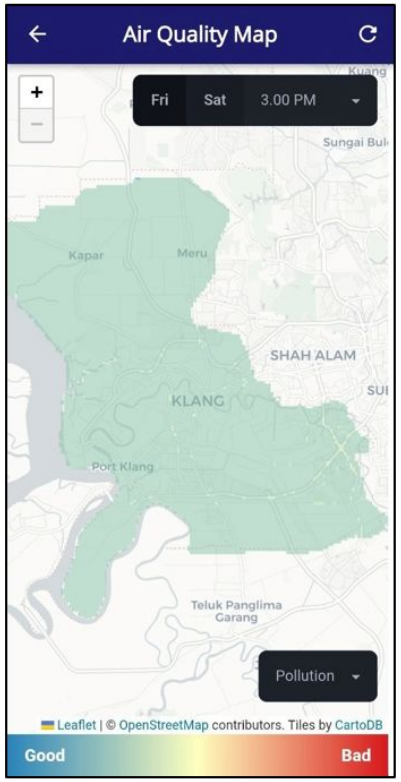
# Research has become a digital and data-intensive enterprise

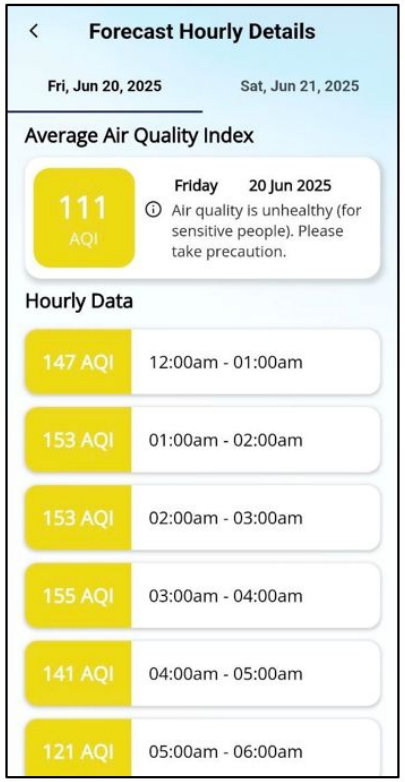
The RESPIRE case study: Leveraging on Environmental data and AI prediction for Solving Healthcare Problem







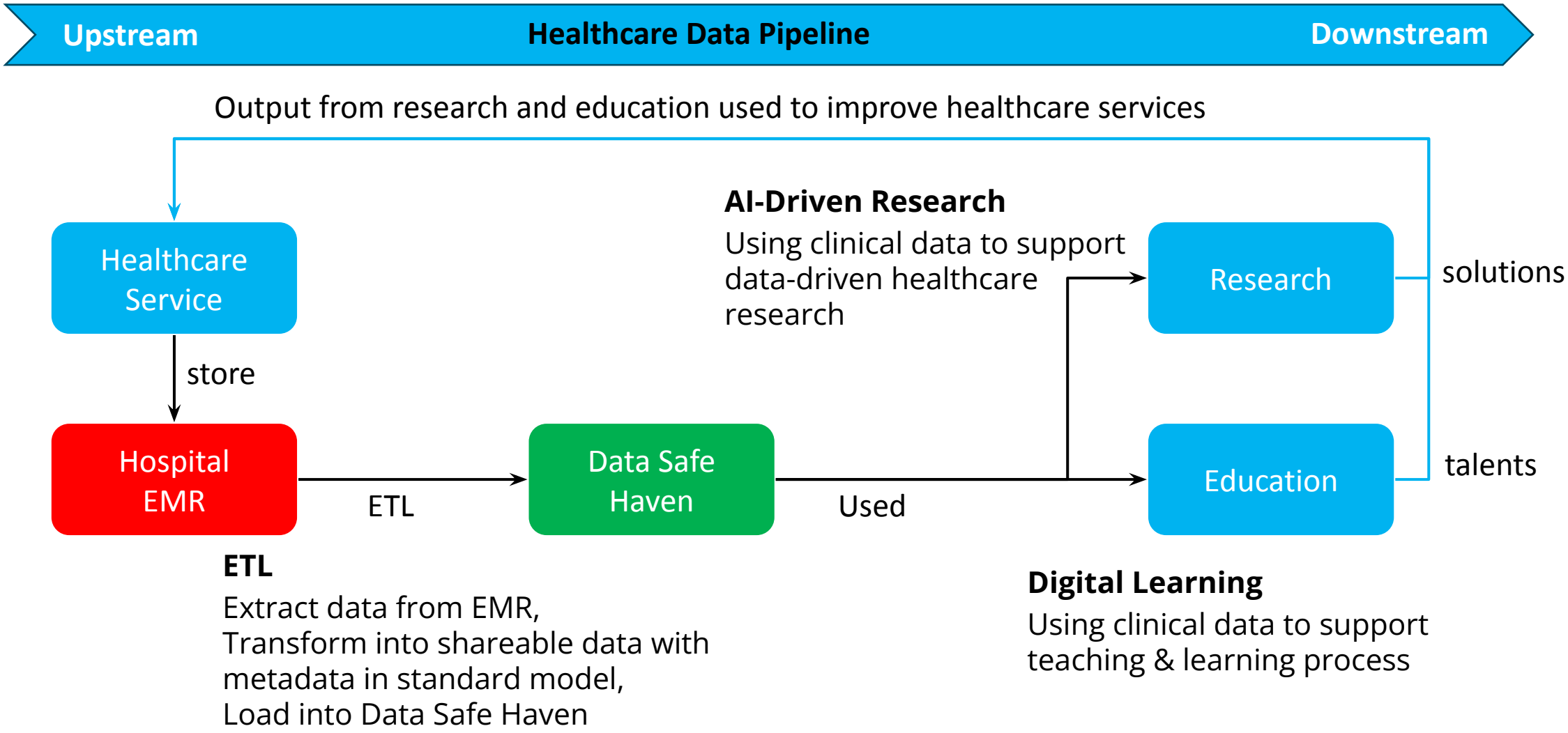



Source: RESPIRE Air Quality App (AQA) for asthma self-management – University of Edinburgh, UM, UKM & CERC

# Digital Infrastructure for Healthcare Ecosystem

Creating a safe haven for healthcare data pipeline



cancer research malaysia

# Mobile Mouth Screening Anywhere®




* All slides and content is proprietary information of Cancer Research Malaysia and are not to be reproduced, distributed, or edited without expressed permission of CRMY.

# The Problem : A Matter of Time & Accessibility



**74%**

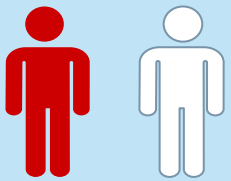
**Oral Cancer Deaths in Asia**



**Late Detection**

**3 of 4** oral cancer patients are diagnosed at **late stage**

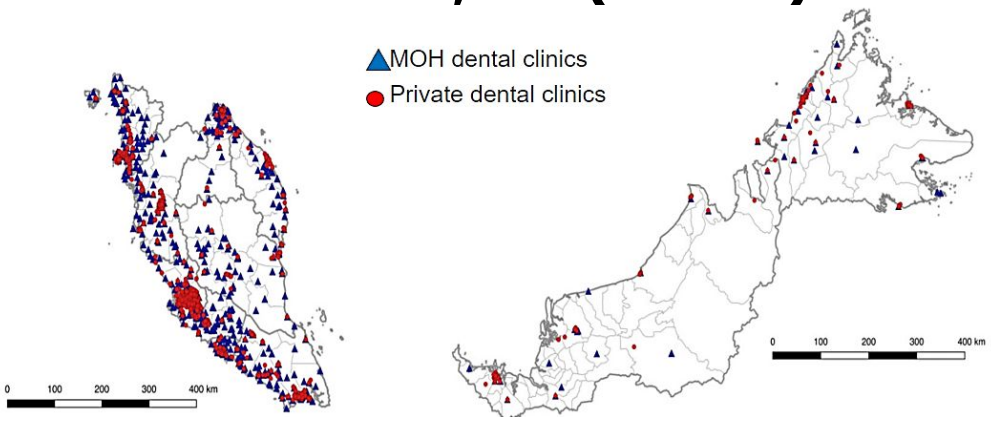




**1 in 2**

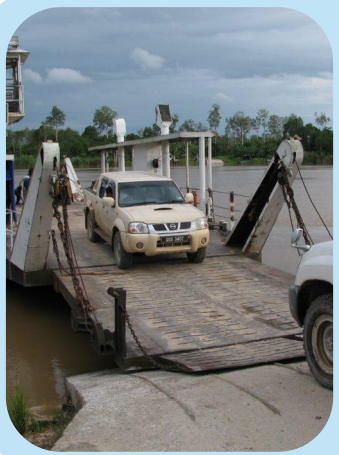
Malaysians with Oral Cancer **die within 2 years**

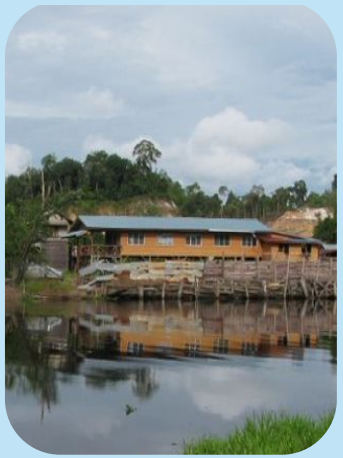
**Public Dental Clinic : Population**
**1:83,000 (Sarawak)**
**1:89,000 (Sabah)**






**Limited Accessibility to Specialist**
(only available at centralized tertiary centres)









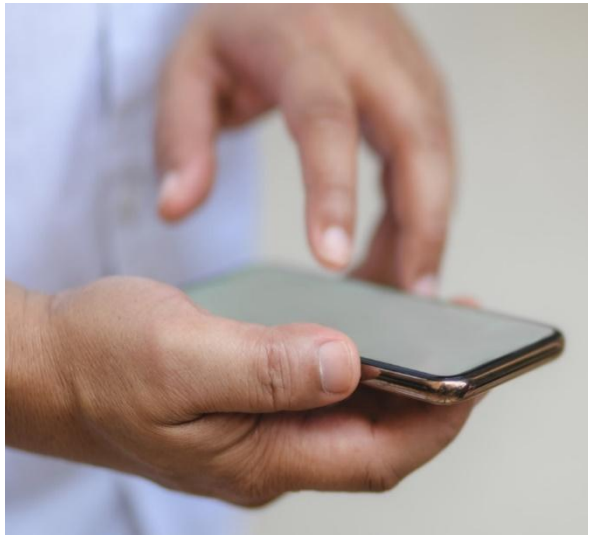
**Geographical Barriers**



* All slides and content is proprietary information of Cancer Research Malaysia (CRMY) and are not to be reproduced, distributed, or edited without expressed permission of CRMY.

# The Solution: Leveraging Widespread Mobile Technology & Photographability of Oral Lesions

**94%**
**Access to Smartphones in Rural Areas**



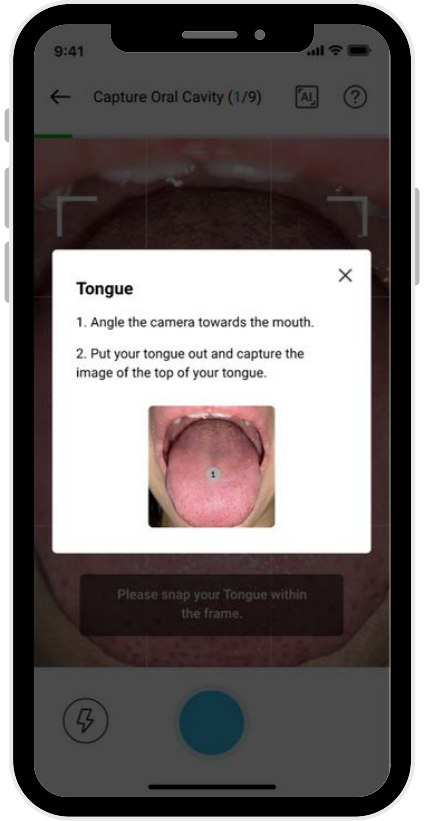
Ref: ICT Use and Access by Individuals and Households 2023 Survey Report

**Oral Cancer Signs**
**(Visible + Photographable)**





**MeMoSA®:**
**Remote Oral Cancer Screening**






* All slides and content is proprietary information of Cancer Research Malaysia (CRMY) and are not to be reproduced, distributed, or edited without expressed permission of CRMY.

# Evolution of MeMoSA: From Prototype to the Real World

cancer research malaysia

## 2017

First prototype developed




## 2018

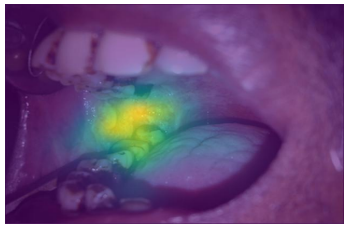
Used prototype to demonstrate that teleconsultation has a **94.9%** accuracy compared to the gold standard of clinical oral examination (COE) [1]



## 2020

MeMoSA used clinically at UM during the COVID-19 pandemic [2]

Proof-of-concept AI models developed [3,4]



## 2023

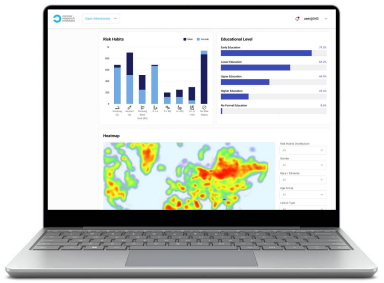
New version of app developed ready for AI integration

New app pilot tested with Healthcare Volunteers (HCVs) in Sarawak [5]



## 2024

MeMoSA Data Management Platform (Workbench) developed




**References:**

1. Haron N, Rajendran S, Kallarakkal TG, et al. High referral accuracy for oral cancers and oral potentially malignant disorders using telemedicine. Oral Dis. 2023;29(2):380-389
2. Rajendran S, Mohamad Zaini Z, Lim YZ, et al. Accelerated implementation of teleconsultation services for the monitoring of oral potentially malignant disorders as a result of the COVID-19 pandemic. Oral Oncol. 2022;126:105730
3. Welikala, R.A. et al. (2020). Fine-Tuning Deep Learning Architectures for Early Detection of Oral Cancer. Mathematical and Computational Oncology. ISMCO 2020. Lecture Notes in Computer Science(), vol 12508. Springer, Cham.
4. Welikala, R.A. et al. (2020). Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer. IEEE Access, vol. 8, pp. 132677-132693.
5. [in publication]



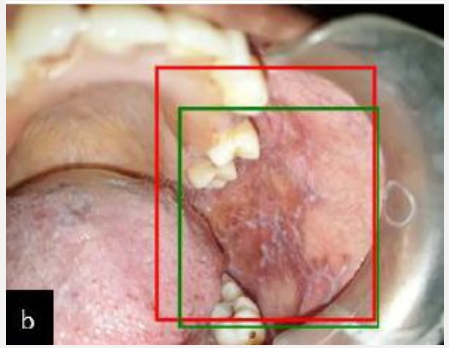
* All slides and content is proprietary information of Cancer Research Malaysia (CRMY) and are not to be reproduced, distributed, or edited without expressed permission of CRMY.

# MeMoSA AI: Automating Lesion Identification

cancer research malaysia

## Published AI Work

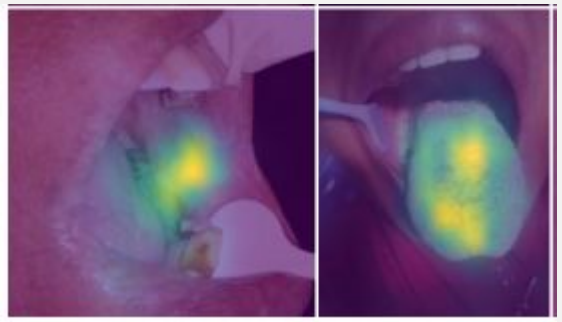
### Deep learning-based image classification[1,2]




### Mouth landmark detection guidance[3]



### Clinically guided soft attention mechanisms[4]



## MeMoSA Dataset[5]: The Largest Known Multi-Country Dataset of Oral Images




**30,039** images from

**5** Asian Countries

**6,426** individuals


**References:**

1. Welikala, R.A. et al. (2020). Fine-Tuning Deep Learning Architectures for Early Detection of Oral Cancer. Mathematical and Computational Oncology. ISMCO 2020. Lecture Notes in Computer Science(), vol 12508. Springer, Cham.
2. Welikala, R.A. et al. (2020). Automated Detection and Classification of Oral Lesions Using Deep Learning for Early Detection of Oral Cancer. IEEE Access, vol. 8, pp. 132677-132693.
3. Lim, J.H. et al. (2021). D'OraCa: Deep Learning-Based Classification of Oral Lesions with Mouth Landmark Guidance for Early Detection of Oral Cancer. Medical Image Understanding and Analysis. MIUA 2021. Lecture Notes in Computer Science(), vol 12722. Springer, Cham.
4. Welikala, R.A. et al. (2021). Clinically Guided Trainable Soft Attention for Early Detection of Oral Cancer. CAIP 2021. Lecture Notes in Computer Science(), vol 13052. Springer, Cham.
5. Chew, S., Nabil, A., Sin, W.J. et al. MeMoSA dataset: A multi-country collection of over 30,000 oral mucosa images with clinically labelled lesions. Sci Data (2026).



* All slides and content is proprietary information of Cancer Research Malaysia (CRMY) and are not to be reproduced, distributed, or edited without expressed permission of CRMY.

# Real World Impact: MeMoSA®
# Teleconsultation Initiative at Universiti Malaya




## MeMoSA® Teleconsultation Workflow:

Patients upload image for specialist review from home.

Digital consultation and recommendations.

Continuous remote monitoring.



## MeMoSA® resolved issues during the COVID-19 pandemic:

- Disrupted Patient Follow-Ups
- Increased Risk of Late-Stage Diagnoses
- Limited Access to Specialists due shortage of in-person consultations

## Results & Patient Responses

**83%** of patients felt MeMoSA® enhanced their access to care

**98%** of patients reported significant time savings

**100%** of participants expressed interest in continuing teleconsultation

**70%** of patients wanted ongoing teleconsultation integration into their routine care



* All slides and content is proprietary information of Cancer Research Malaysia (CRMY) and are not to be reproduced, distributed, or edited without expressed permission of CRMY.

# Real World Impact: Closing the Gap in Rural Sarawak

cancer research malaysia

## MeMoSA Workflow with Healthcare Volunteers (HCVs)




**Village community**

**HCVs**
Capture images and clinical information on MeMoSA App

**MeMoSA**

**Primary care dentist**
Reviews images on app and decides if further referral is needed

**MeMoSA**

**Specialist**
Reviews referred images on app and provides advice on management and final referral decision







**47** Trained HCVs

**66** Trained Clinicians

# Real World Impact: Our Impact in Sarawak

cancer research malaysia

**Overall Reach** **87.7%**

**1863**

**people screened over 1 year (10x more than 2022)**







**28**

**unnecessary trips to the specialist prevented**

**3**

**patients with oral potentially malignant lesions (OPMD) identified and referred for treatment**

**Median time taken by dentists to review cases via MeMoSA®: ahead of the 48-hour target**

2 HOURS

**70%**

**of images captured by Healthcare Volunteers (HCVs) via MeMoSA® were good quality**




cancer research malaysia

## Contact us:

*Prof. Dr Cheong Sok Ching, FASc*
sokching.cheong@cancerresearch.my
memosa@cancerresearch.my

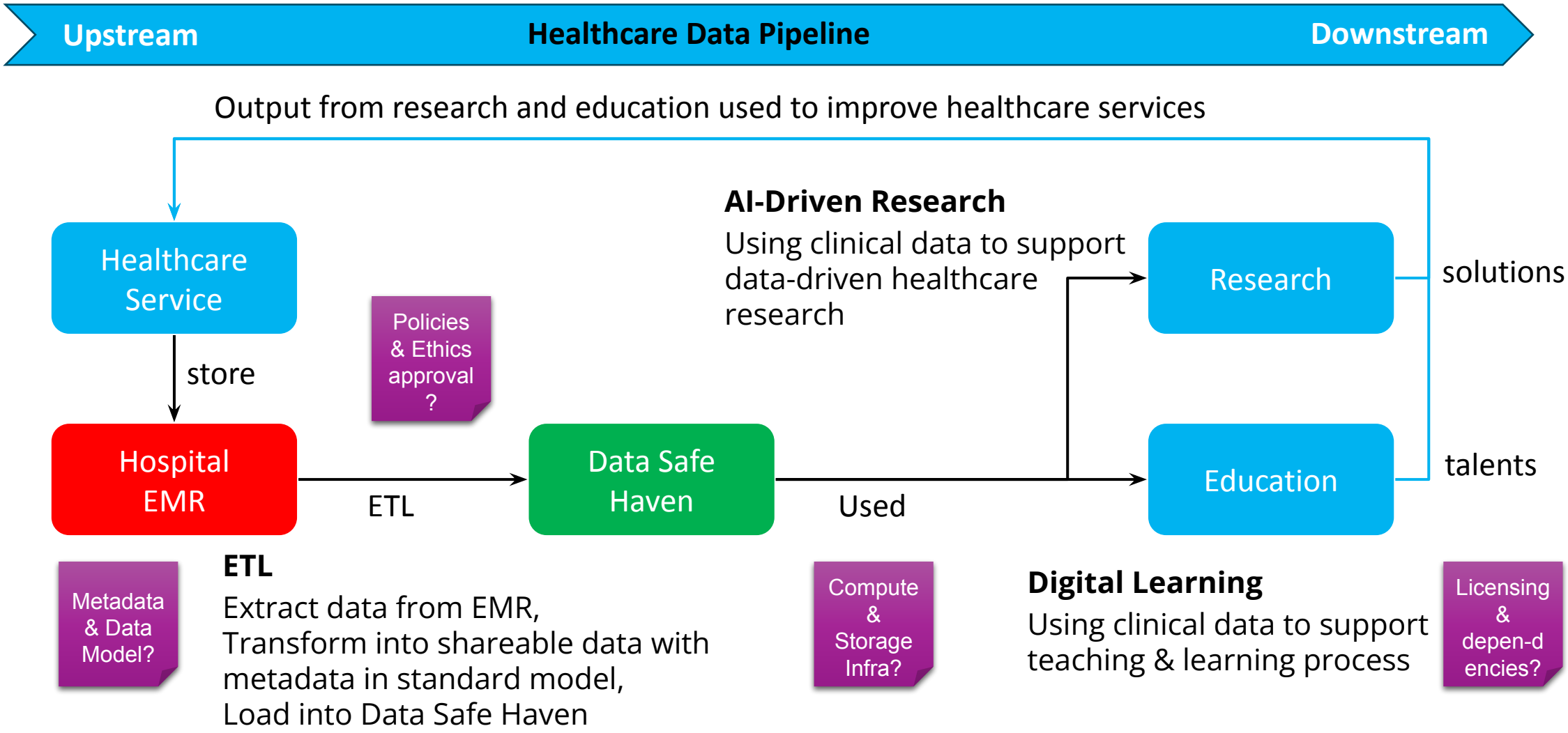
Cancer Research Malaysia
(200001007481)
Level 1, Subang Jaya Medical Centre South Tower,
No. 1, Jalan SS 12/1A,
47500 Subang Jaya,
Selangor, Malaysia.



* All slides and content is proprietary information of Cancer Research Malaysia (CRMY) and are not to be reproduced, distributed, or edited without expressed permission of CRMY.

# Digital Infrastructure for Healthcare Ecosystem

Creating a safe haven for healthcare data pipeline

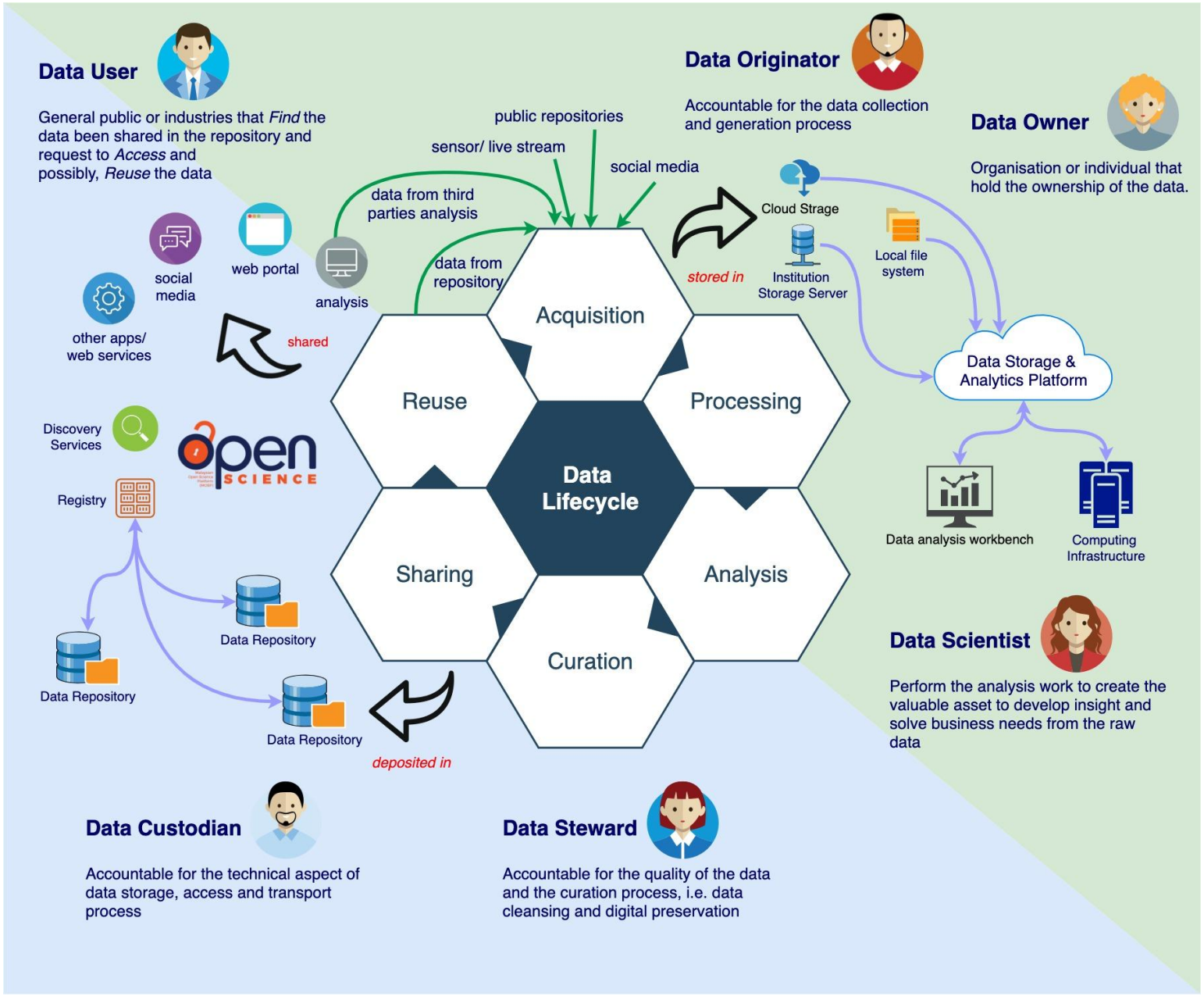


Data Services

Data Registry

Data Aggregator

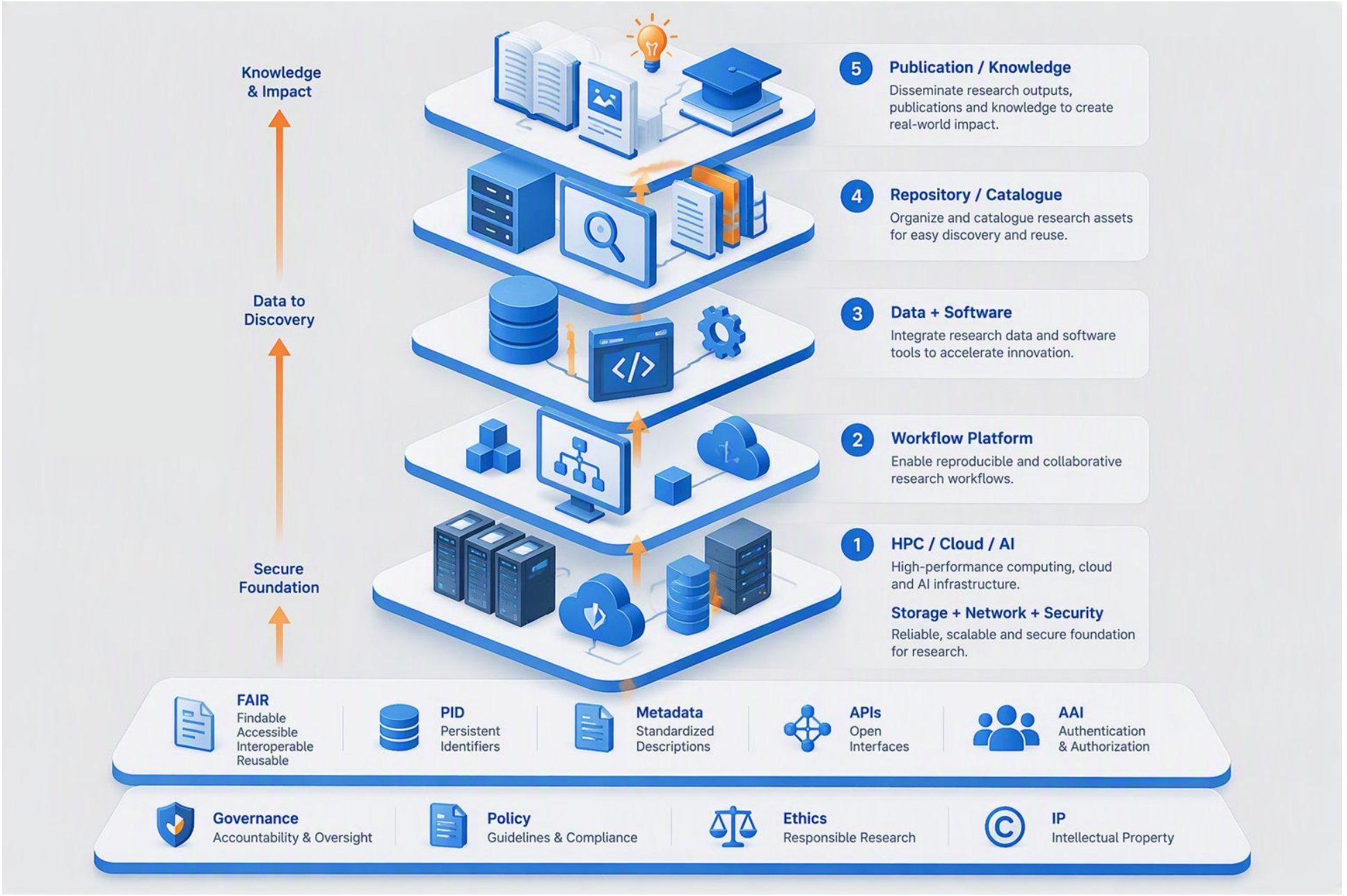
Data Repository




Data Storage System

Data Computing Platform

Data Curation Platform




# *Modern Research Ecosystem*

# European Open Science Cloud

- **Resource federation**
  - Rather than creating one enormous European repository, the model is to **connect** research infrastructures, services, data and computing resources through common standards and governance.
- The EOSC Partnership identifies three major objectives:
  - make Open Science practices and skills normal and rewarded
  - enable standards, tools and services to find/access/reuse research
  - establish sustainable federated infrastructure for scientific results.

*Connect existing infrastructures rather than duplicate them*

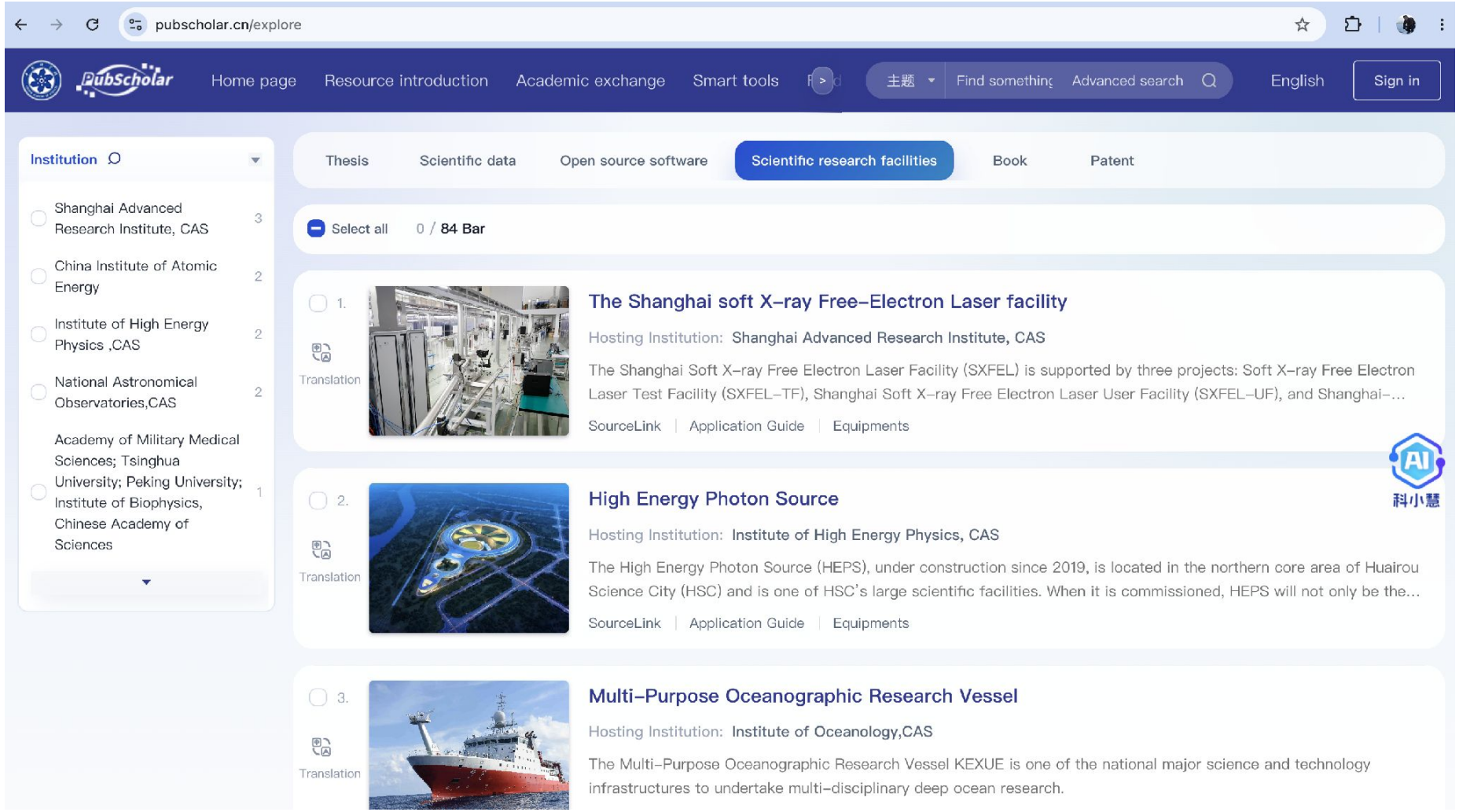


# China: infrastructure at national scale

- The CAS' Computer Network Information Center (CNIC) operates **infrastructure spanning data, cloud, HPC, AI computing**, identity/trust frameworks and open research infrastructure, and the CAS' National Science Library builds and operates the PubScholar Public Academic Platform.
- Open for Science10 (OS10) initiative, endorsed by UNESCO in 2025, aims to develop a **globally connected** and **interoperable scientific infrastructure network**.

*Integrated IaaS/PaaS/SaaS layers supporting massive data transfer, cross-domain scheduling, and AI-driven analysis.*

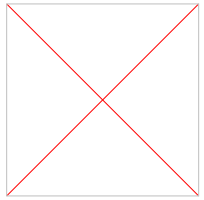
# Openness requires interoperability and shared access




*not necessarily one physical location*

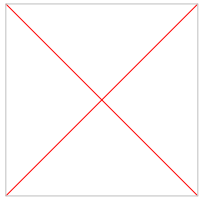
# How does Open Science Empower AI?

Open Science provides the data, knowledge and collaborative ecosystem that AI needs to advance research



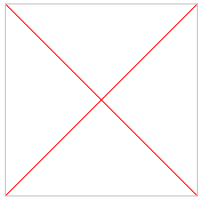
## Talent

The people to drive the innovation and make the change



## Technology

The infrastructure including hardware, software and services as the enabler for AI adoption



## Data

AI needs good data, i.e. quality and quantity + legally usable

*" AI is only as useful as the research ecosystem behind it "*

# Open Science makes AI research more reproducible

*Publishing an AI paper alone is insufficient to reproduce its findings.*

*Researchers also need appropriate access to training and evaluation data, code, model versions, parameters, evaluation methods and computing environments.*



# AI makes scientific knowledge more discoverable

- **Semantic Search Engines**:
  - Platforms like Google Scholar Labs use natural language processing to understand the intent and context behind complex research questions.
- **Cross-Discipline Knowledge Graphs**:
  - AI connects scattered clues across entirely different fields. For example, AI workflows have successfully linked raw biological experimental data with decades of separate, unstructured literature to uncover hidden mechanisms behind neurodegenerative diseases.
- **Breaking Language Barriers**:
  - AI tools provide contextual translation for non-native English speakers. This makes localized or regionally restricted open data globally discoverable, balancing global publication disparities.

# AI automates research workflows

- AI can help with data cleaning,
  code generation, metadata enrichment, anomaly detection, literature screening and workflow orchestration.
- Use case:
  - In a multi-institution climate study, AI-assisted tools could help standardise incoming metadata, flag inconsistent observations and prepare datasets for a documented analysis workflow.
  - **Researchers still validate transformations and scientific conclusions.**

# Five Practical Principles




## 1 Open by design, not as an afterthought

Build data management, documentation, persistent identifiers and responsible sharing into the research lifecycle from the beginning

## 2 Interoperability before duplication

Connect existing repositories, software, computing services and institutional systems through shared standards and interfaces.

## 3 Responsible openness and trustworthy AI

Respect privacy, consent, intellectual property, sensitive information and indigenous knowledge. Make AI outputs traceable, testable and subject to human oversight.

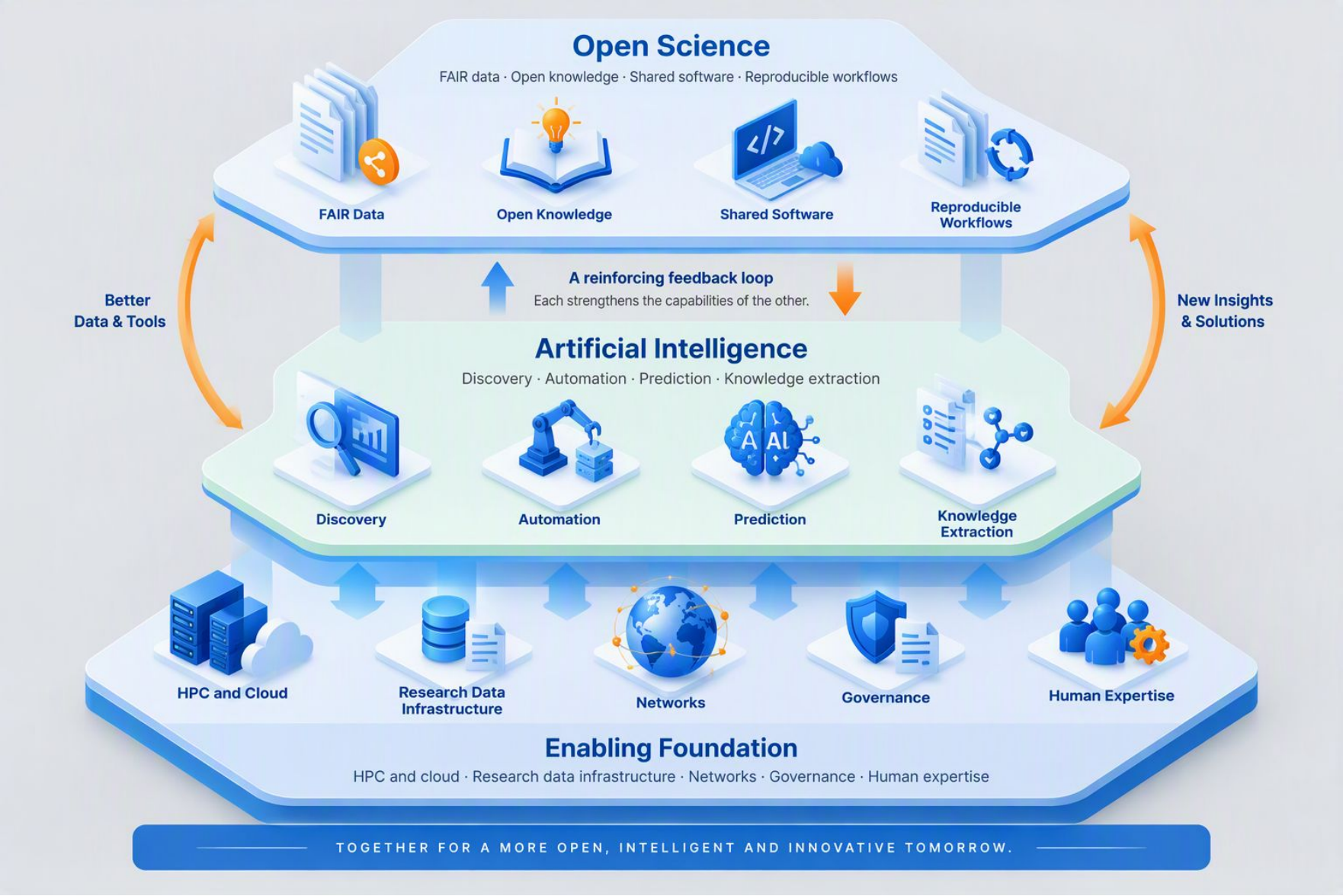
## 4 Reproducibility throughout the AI lifecycle

Preserve the provenance of data, code, models, parameters and computational workflows so that results can be checked and, where possible, reproduced.

## 5 Treat infrastructure as a long-term public research capability

Invest in people, stewardship, sustainable services and shared computing—not just hardware purchases or short-term projects.

# A two-way relationship



# From digital infrastructure to intelligent Open Science

- Data without interoperability limits discovery.
- AI without trustworthy scientific foundations risks amplifying errors.
- Open Science and AI together can accelerate discovery—but only when we invest in sustainable, inclusive and responsible infrastructure.

*" Stop thinking of Open Science as simply making research outputs available, or AI as simply deploying intelligent models…*
*Let's build an* ***interconnected research ecosystem*** *"*

# Thank you

sunwayuniversity.edu.my

Chee Sun LIEW cheesunl@sunway.edu.my

