Data driven healthcare in the South Africa

By Cara Pienaar

Data Analyst at Jembi Health Systems

Consulting for JHS at the Western Cape Provincial Health Data Centre

cara.pienaar@jembi.org

About this talk

- A data driven approach to all aspects of healthcare
- As opposed to just collecting data
- How it can be implemented in a low resource setting with a complicated history, like South Africa

Specifically in the Western Cape Province

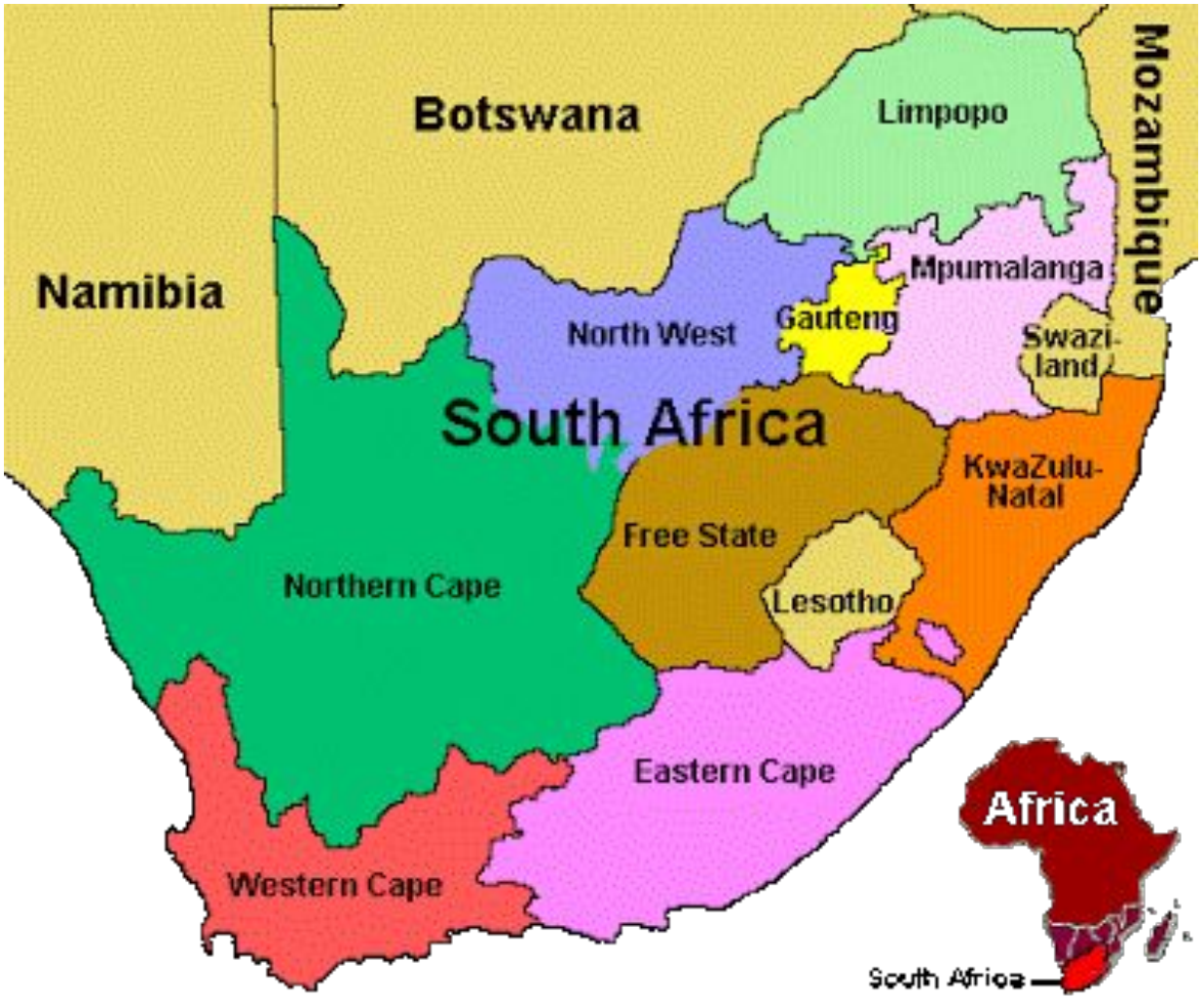


Image sources: 2

About my job

- Jembi Health Systems is a non-profit that works in Africa to strengthen health information systems
- Like a tech company, but for governments and communities
- Two of our projects are the OpenHIM and the Single Patient Viewer
- Both are used by the Western Cape Provincial Health Data Centre
- I spend half my time Jembi and half my time at the Western Cape Dept of Health

Outline

- History and policy
- Western Cape Data Centre Architecture
- Patient Master Index
- Episodes
- Single Patient Viewer
- Other uses
- Data Governance

Context: a history of fragmentation

The modern healthcare system

- Dutch colonial rule from 1652 - 1800 and British colonial rule from 1800 - 1910.
- By the late 1800s, orthodox Western medicine was widespread, with hospitals in most major centres.
- In 1910, the Union of South Africa was established, with health services divided amongst the provinces.
- In 1919, Provinces were made responsible for curative care and local authorities for preventive (primary) care.

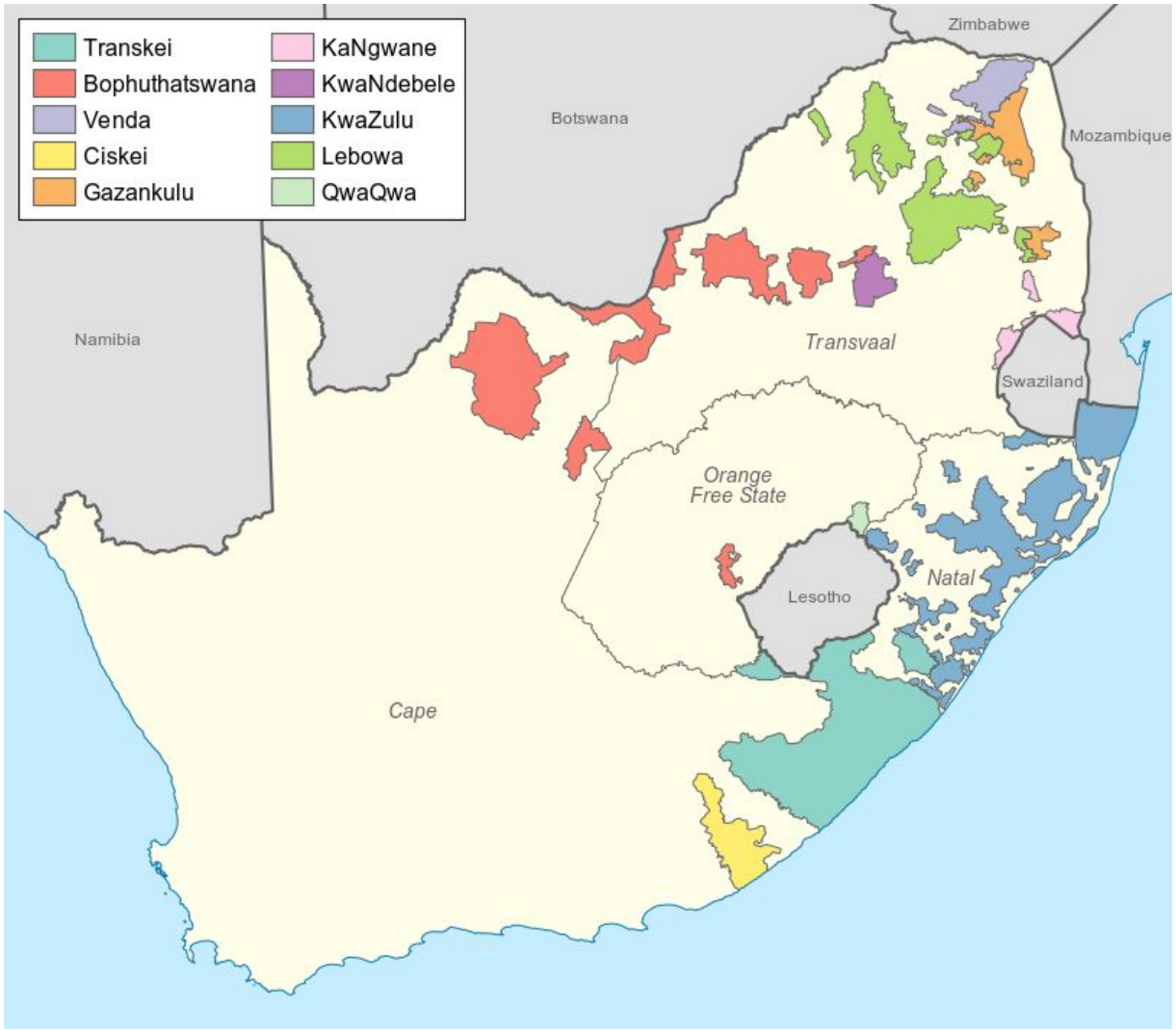
References: 1

The modern healthcare system

Apartheid began in 1948 and further entrenched fragmentation:

- Deliberate rejection of preventive and primary healthcare.
- By 1990, over 60% of doctors worked in the private sector.
- By the end of Apartheid in 1994, 14 separate health departments for white people, coloured people, Indian people, and for each Bantustan (territories set aside for black African people).

References: 1



Map of Apartheid Bantustans.

Image sources: 1

The modern healthcare system

The post-Apartheid government attempted to take a more holistic approach to health.

- In 1996, free primary healthcare was legislated for all.
- Free healthcare to pregnant women and children under 5.
- There was an increased focus on reproductive health, mental health and the restriction of tobacco.
- Attempt to integrate all the disparate systems.

References: 1

Challenges

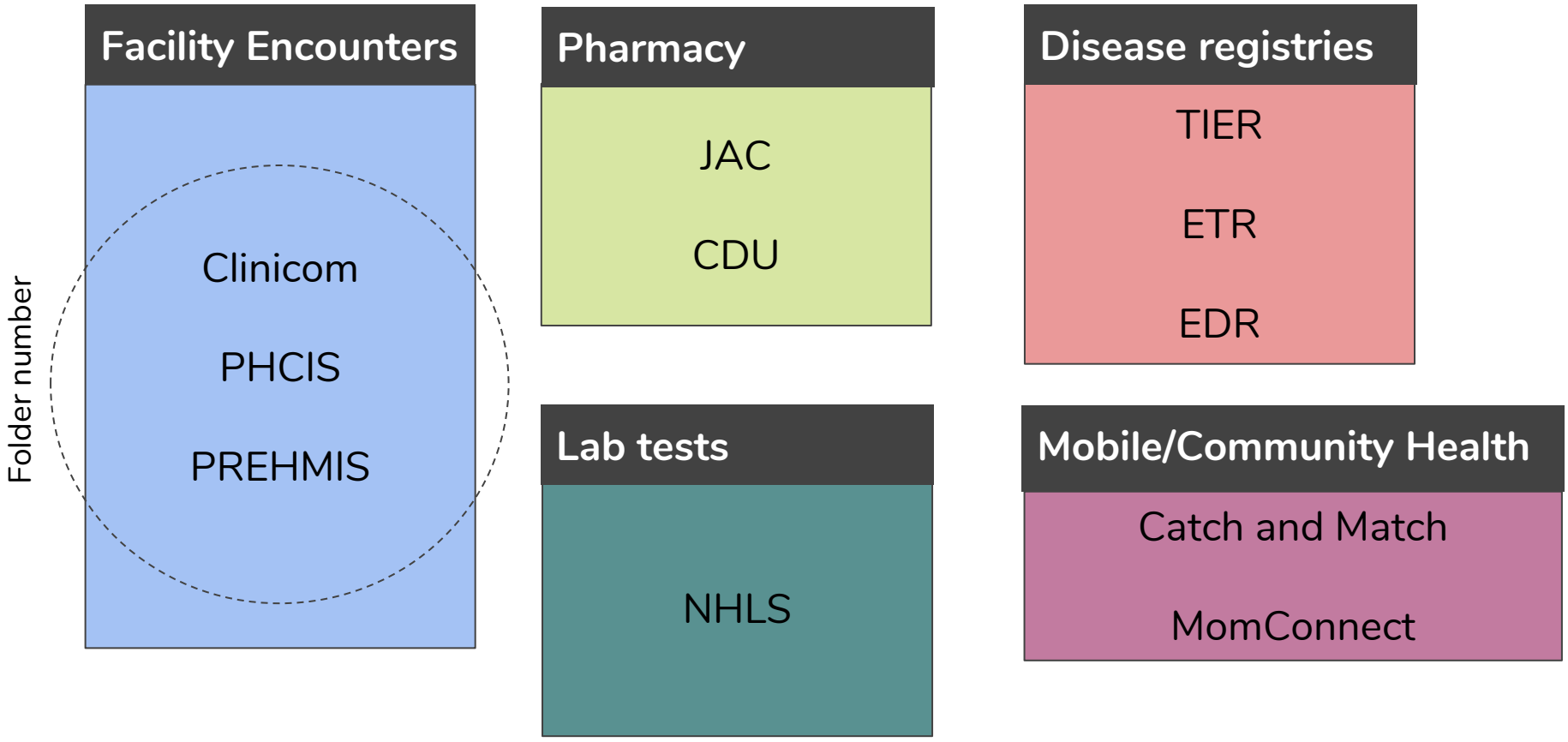
Today we have a variety of impressive digitized systems, storing hospital visits, labs, etc.

However:

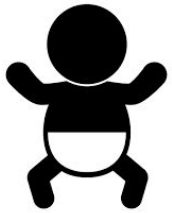
- Many disparate systems with little to no communication or interoperability.
- Inequity of electronic health services across provincial and national departments.
- No single patient identifier.
- Electronic health initiatives do not reach fruition because they can’t integrate.

References: 1,2

Examples of health data sources in the Western Cape



Example: number of births

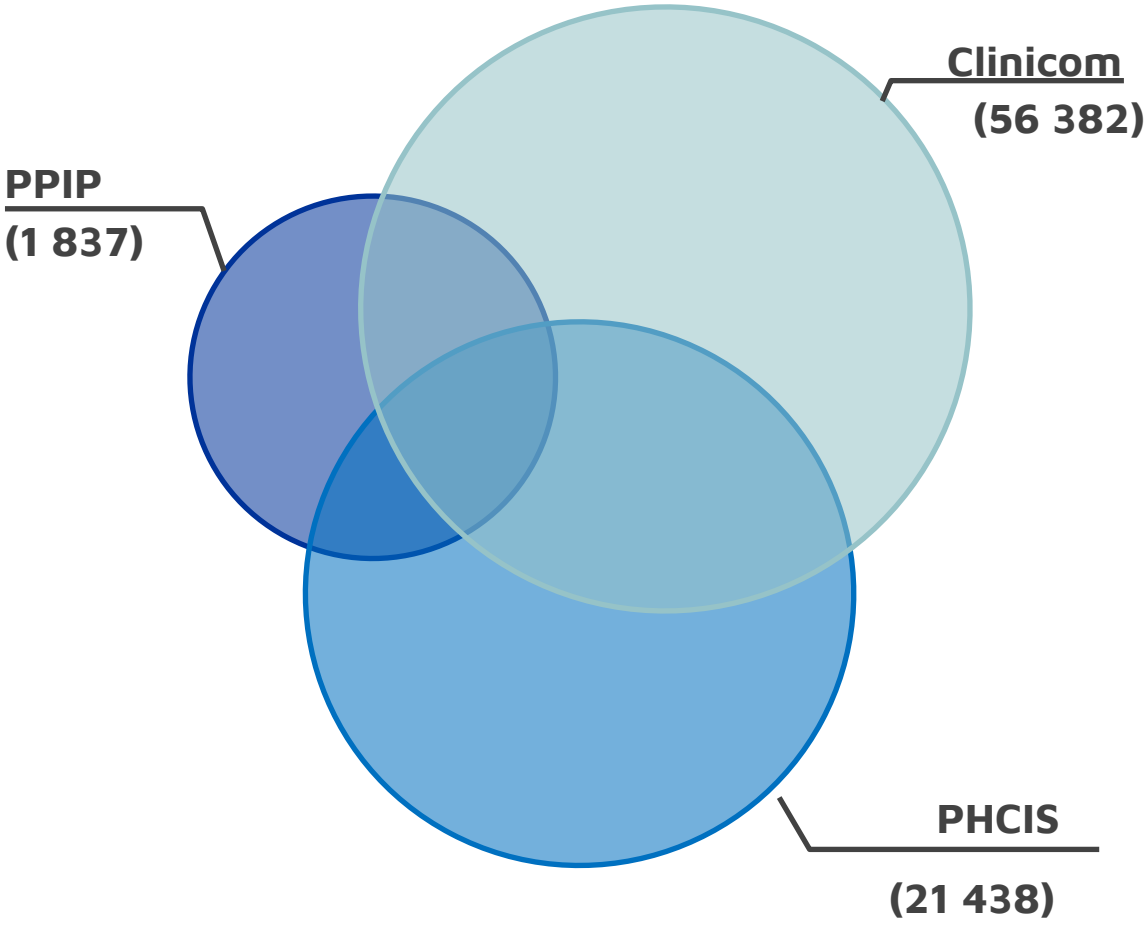


Digital birth data sources in the Western Cape:

- **Primary Health Care Information System (PHCIS)** at clinics
- **Clinicom Patient Management System** at hospitals.
- **Perinatal Problem Identification Program (PPIP)**, a maternal and infant death register at all facilities delivering and caring for newborn babies.

Example: number of births

Overlaps between birth datasources (2015)



References: 3

Data harmonisation

Key words in national policy (2012 and 2014)

eHealth

The use of information technologies including digitized patient records and mobile health tools for:

- treatment,
- research,
- education,
- tracking diseases,
- monitoring public health.

Interoperability

The ability of different information technology systems to:

- communicate,
- exchange data,
- and use the information that has been exchanged.

References: 2,4

Vision for the future

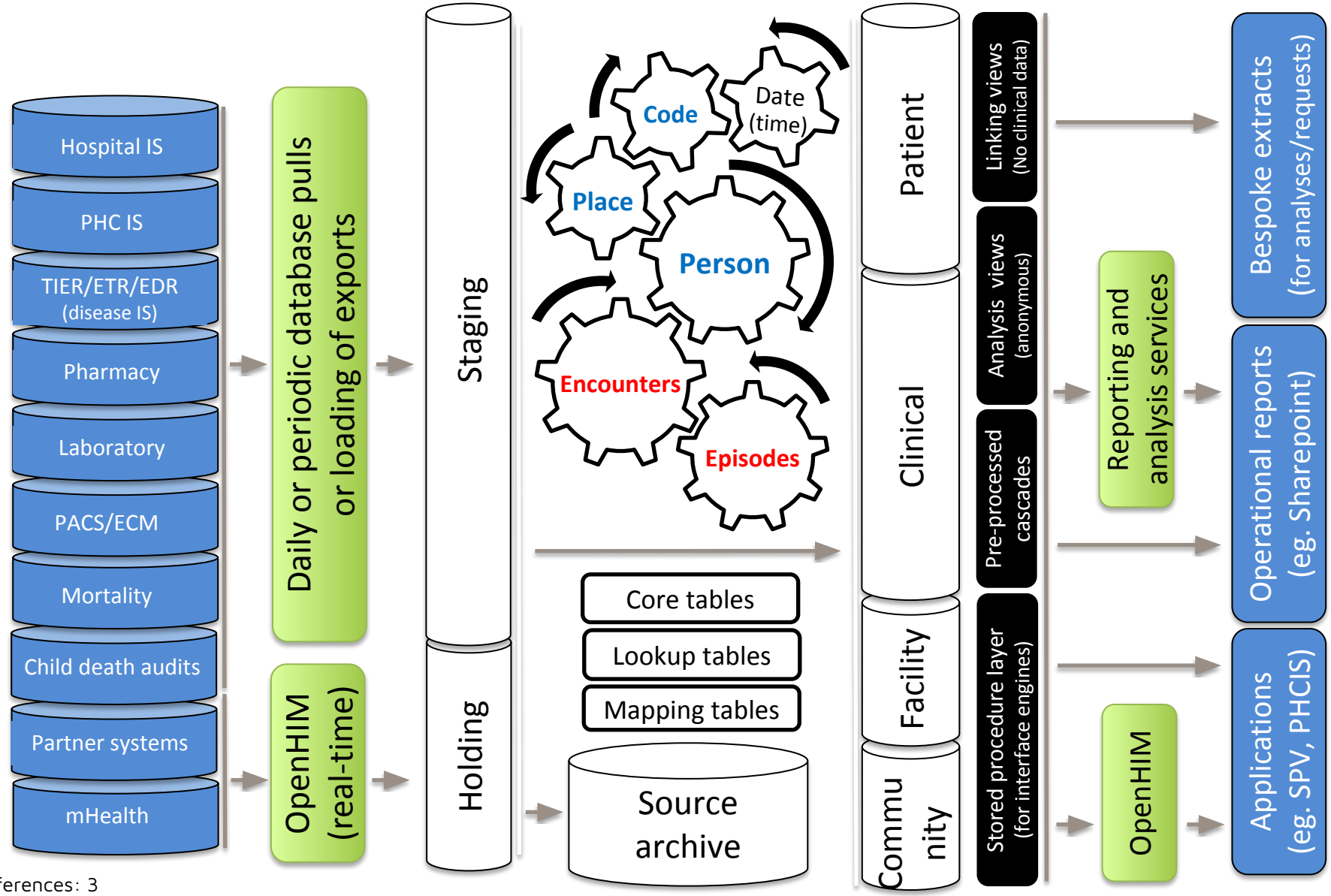
- Develop **eHealth** and **interoperability**
- A **unique patient identifier** for every person
- Central **data warehouses** in each province
- **Complete health records** for every person

Western Cape Provincial Health Data Centre

Western Cape Provincial Health Data Centre

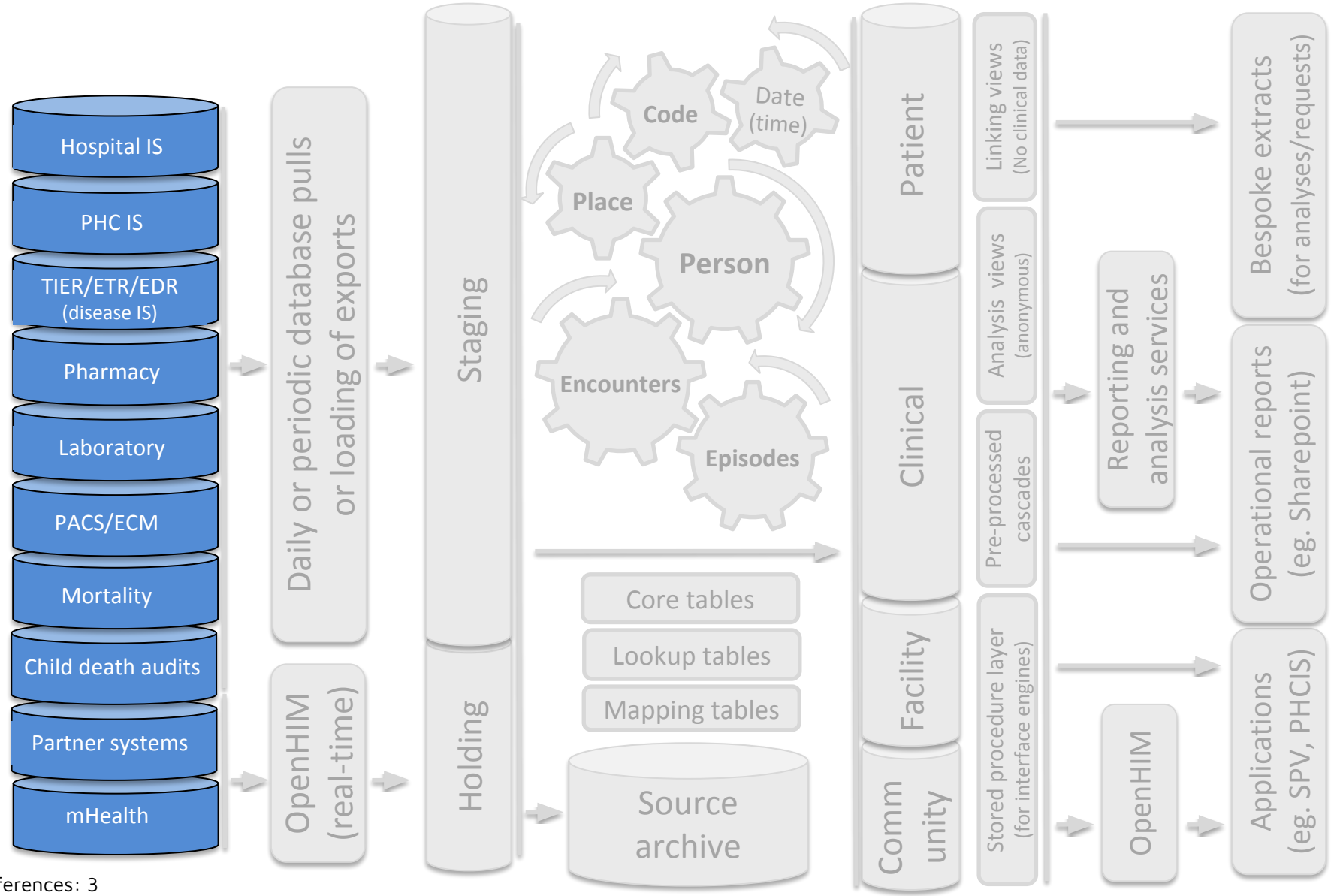
- One implementation of the eHealth strategy.
- Its goal is to be a central repository for all public sector patient data in the Western Cape.
- Brings together clinic visits, hospital admissions, community health worker interactions, lab tests, drugs dispensed and more under a single patient identifier.
- A step towards patient focused care.
- Constantly iterating, prototyping and improving.

High level architecture



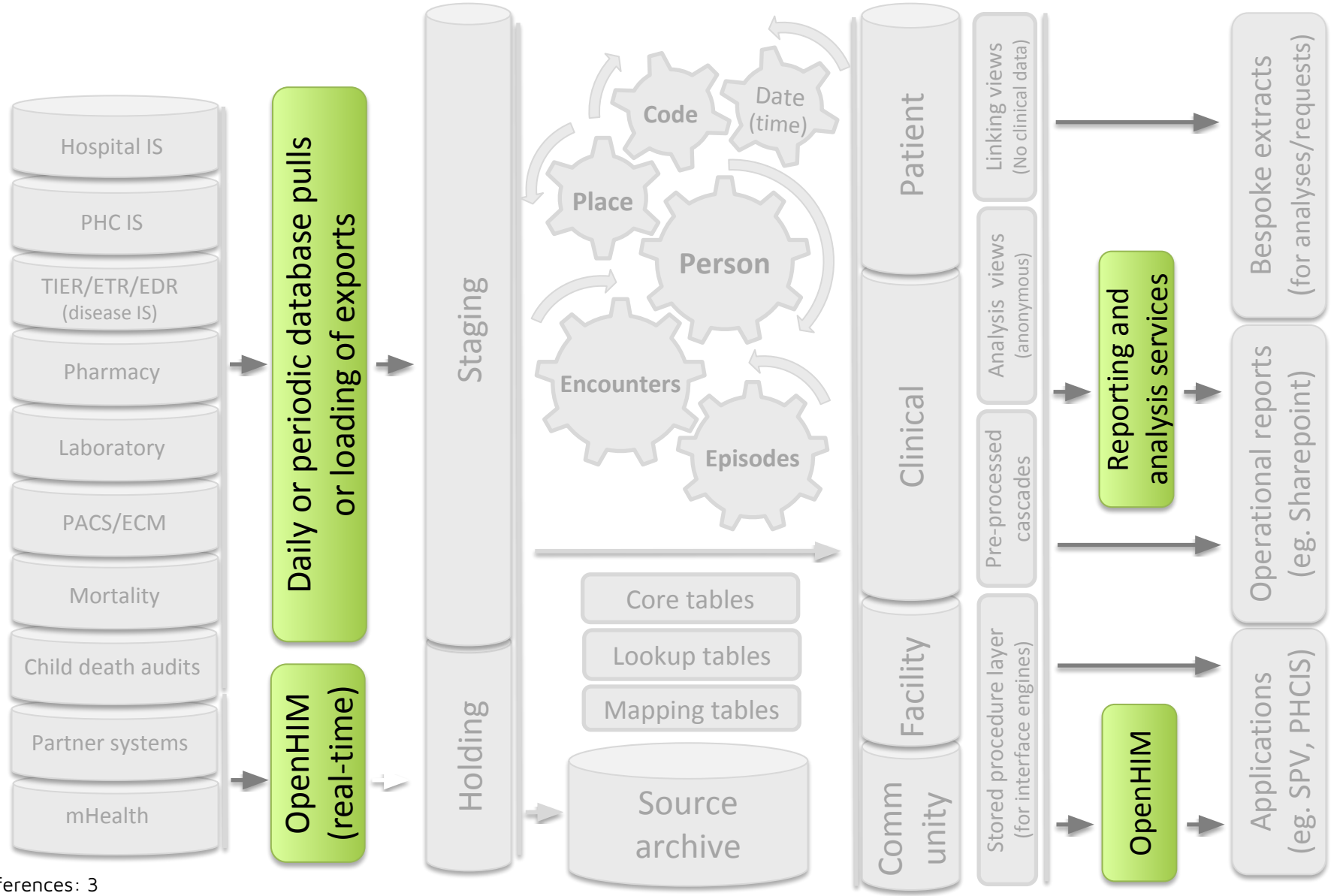
References: 3

Source systems



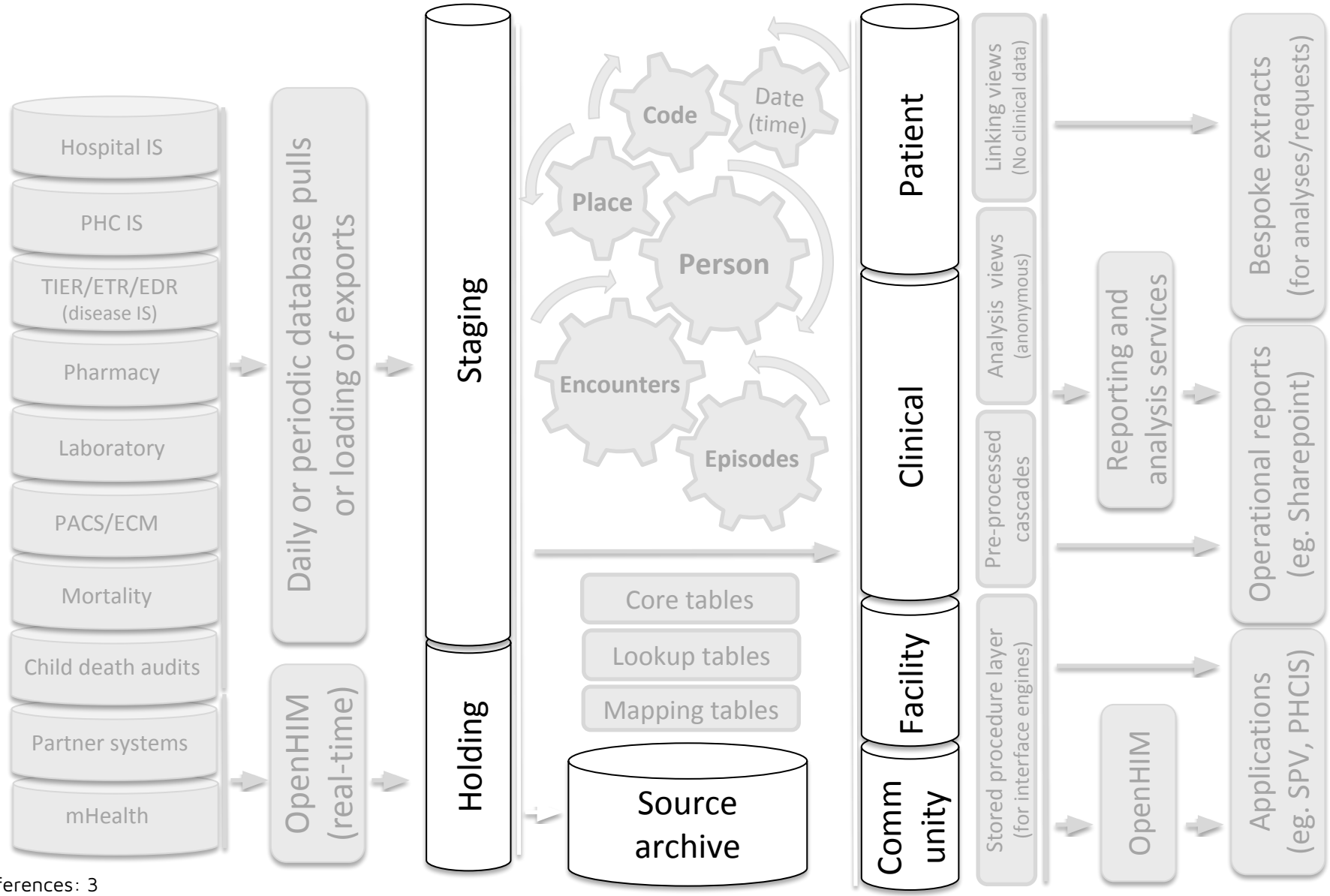
References: 3

Tools for data take-on and access



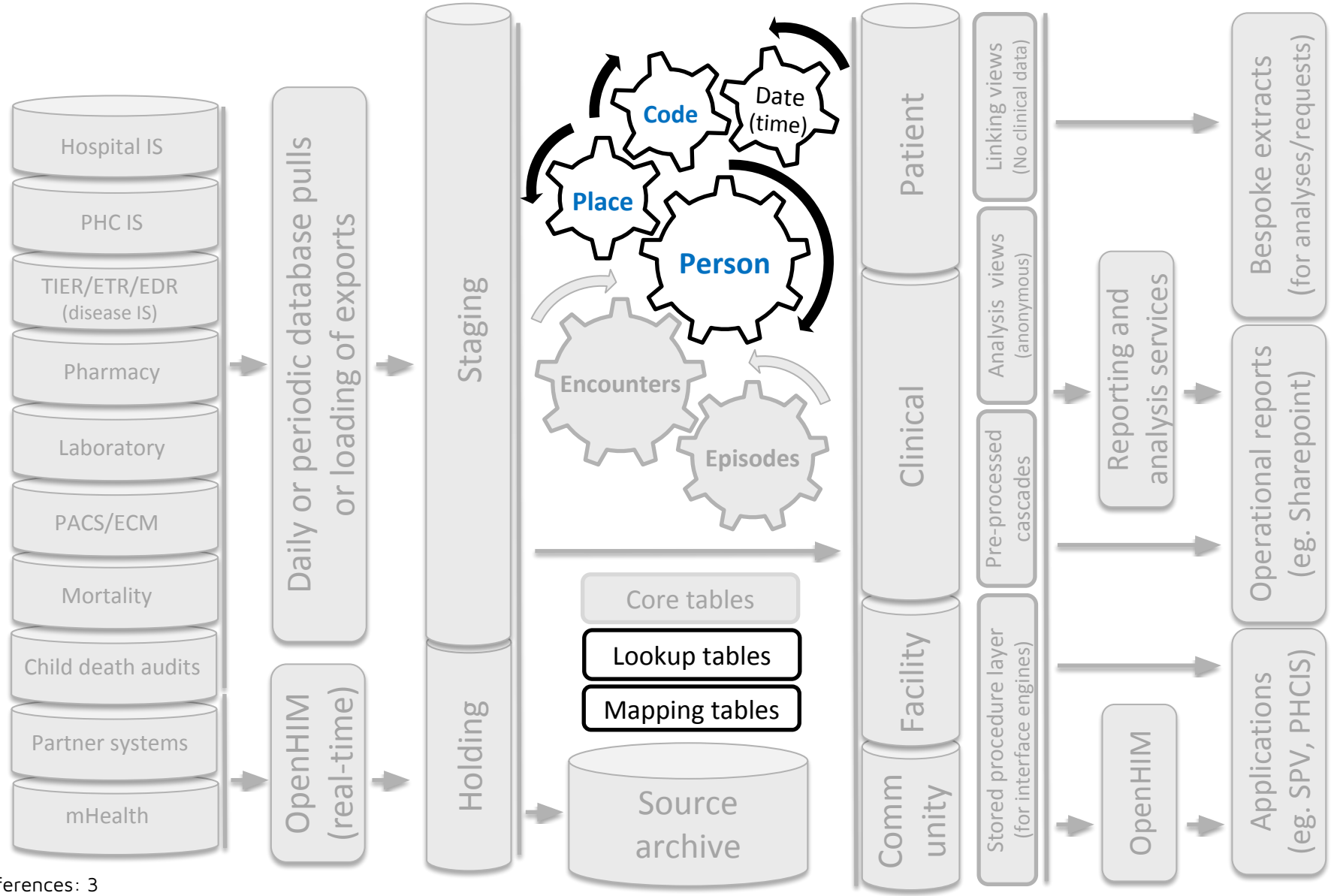
References: 3

PHDC core databases



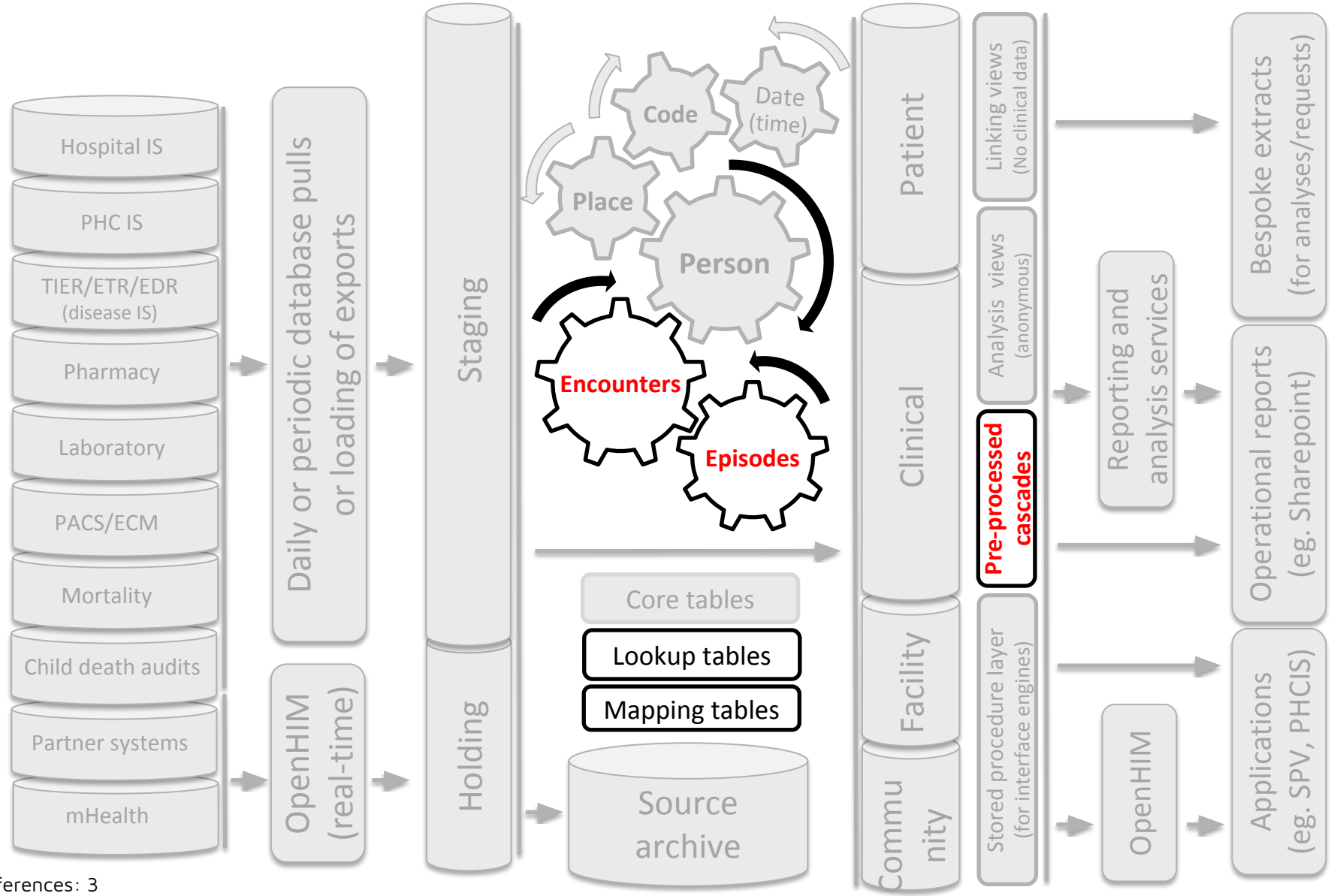
References: 3

Data curation



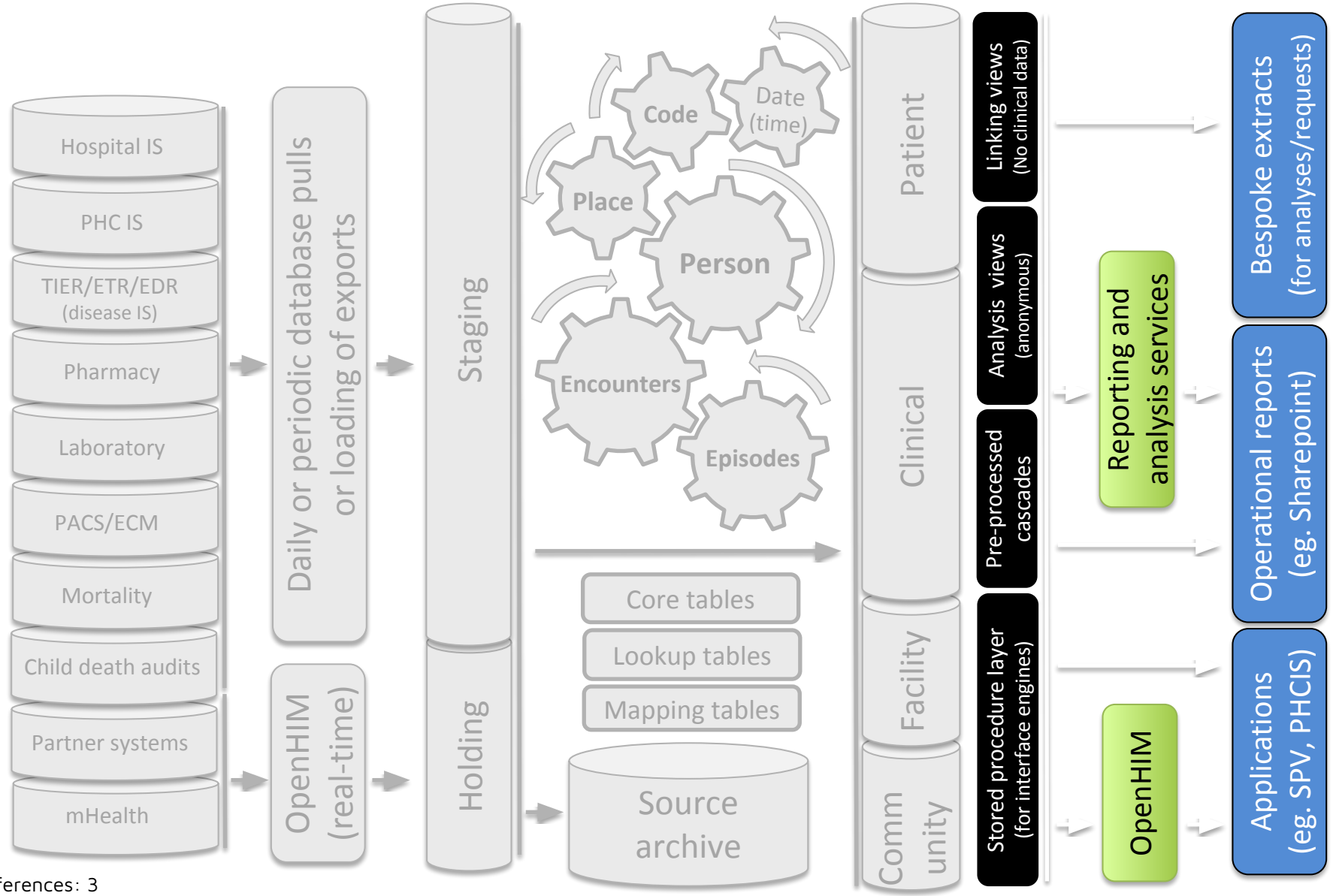
References: 3

Data beneficiation / enrichment



References: 3

Data access and reporting



References: 3

Patient Master Index

Patient Master Index (PMI)

- A big **directory** of patient information attached to an **identifier** called a PMI ID
- Fortunately, the **“folder number”** system was already used throughout Western Cape hospitals and clinics, for over 25 years (used to be paper-based)
- We do a one-to-one mapping from the folder number to the PMI
- PMI is then be linked to other clinical information - it is the folder number *and more*

References: 3

PMI

We store a history of every version of the patient's information ever taken.

PMI ID	Name	Surname	Date of Birth	SA ID Number
523149	Tim	Wolff	1958-02-03	5802031111111
523149	Tom	Wolff-Piggott	1958-02-03	5802031111111
523149	Timothy	Wolff Piggott	1958-02-03	

References: 3

Duplicate PMIs

Sometimes, a patient goes to the hospital and gets issued a new (duplicate) folder number, resulting in a new PMI ID

PMI ID	Name	Surname	Date of Birth	SA ID Number
523149	Tim	Wolff	1958-02-03	5802031111111
981034	Tom	Wolff-Piggott	1958-02-03	5802031111111
523149	Timothy	Wolff Piggott	1958-02-03	

References: 3

Duplicate PMIs

The Data Centre handles this using their own **deterministic algorithm** in the form of a SQL stored procedure.

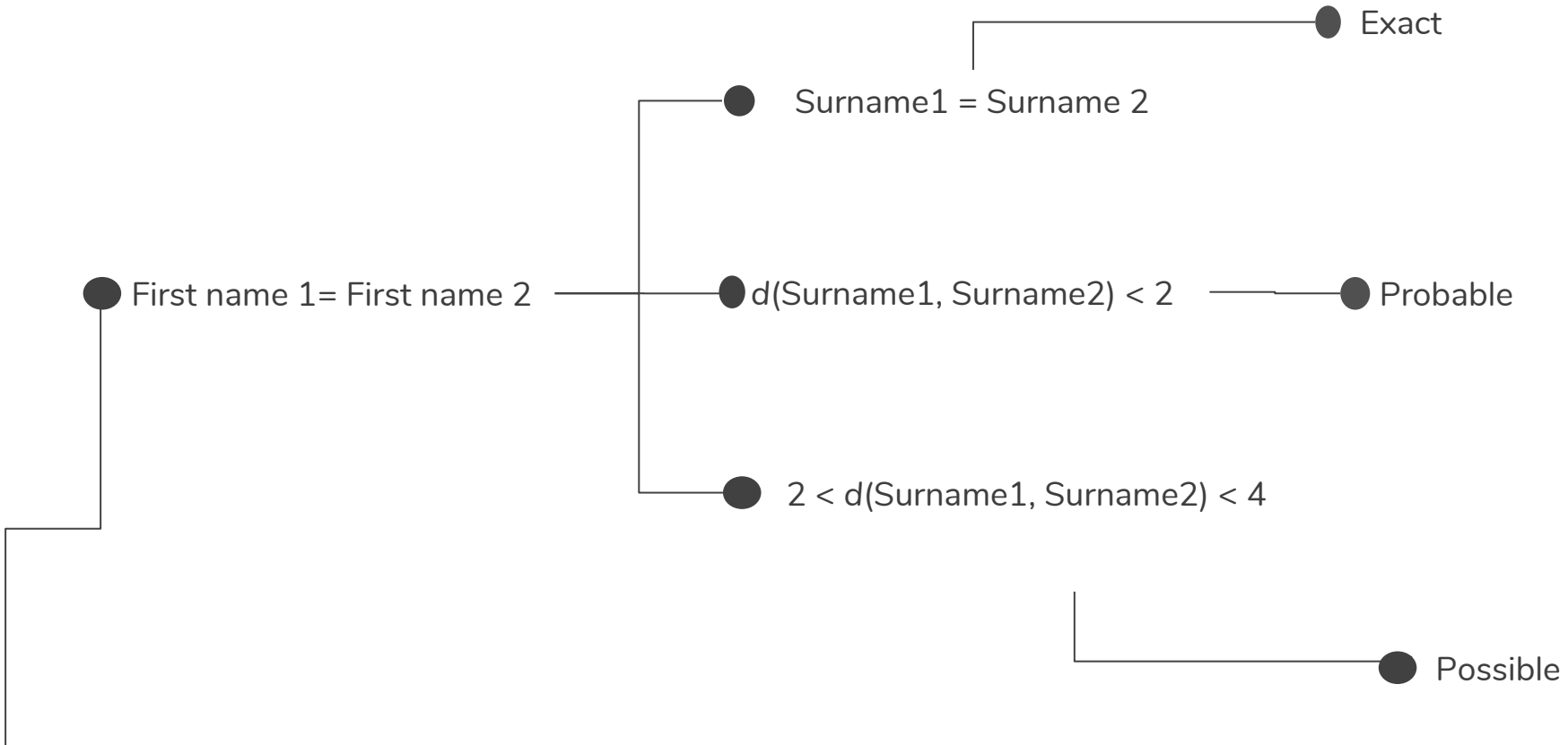
At least three identifiers out of name, surname, dob, address and folder number should match or be close to one another (Levenshtein distance).

Match is then classified as possible, probable or exact.

References: 3

Duplicate PMIs

It’s a giant IF-ELSE statement with nested distance calculations



What if a link is wrong?

It’s ok because it’s just a soft link

Dominant PMI ID	PMI ID	Name	Surname	Date of Birth	SA ID Number
523149	523149	Tim	Wolff	1958-02-03	5802031111111
523149	981034	Tom	Wolff-Piggott	1958-02-03	5802031111111
523149	523149	Timothy	Wolff Piggott	1958-02-03	

- Transparent and visible to doctors

References: 3

Bringing in other data

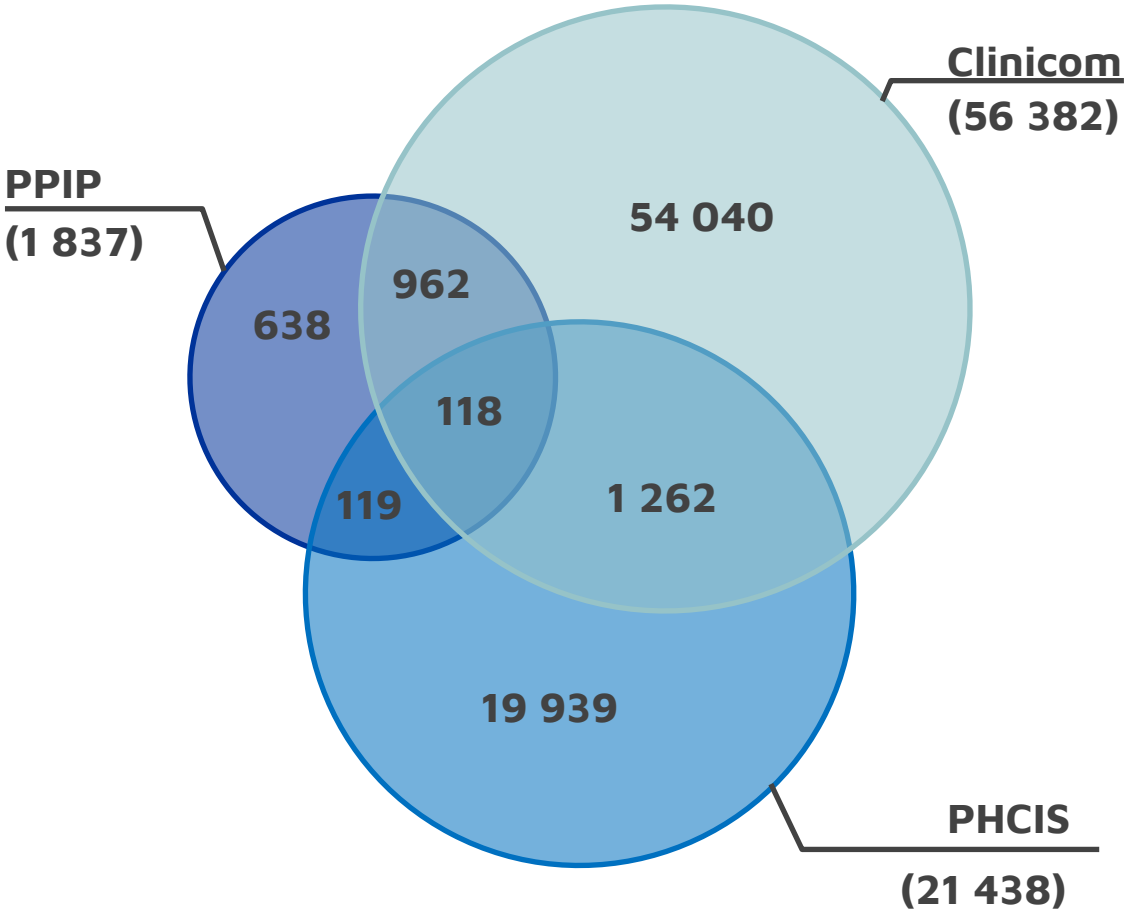
Most data from other sources such as labs and drugs come with patient details such as name and date of birth.

These are matched to the PMI of the patient using a similar deterministic matching algorithm.

References: 3

Example: number of births

Overlaps between birth datasources (2015)



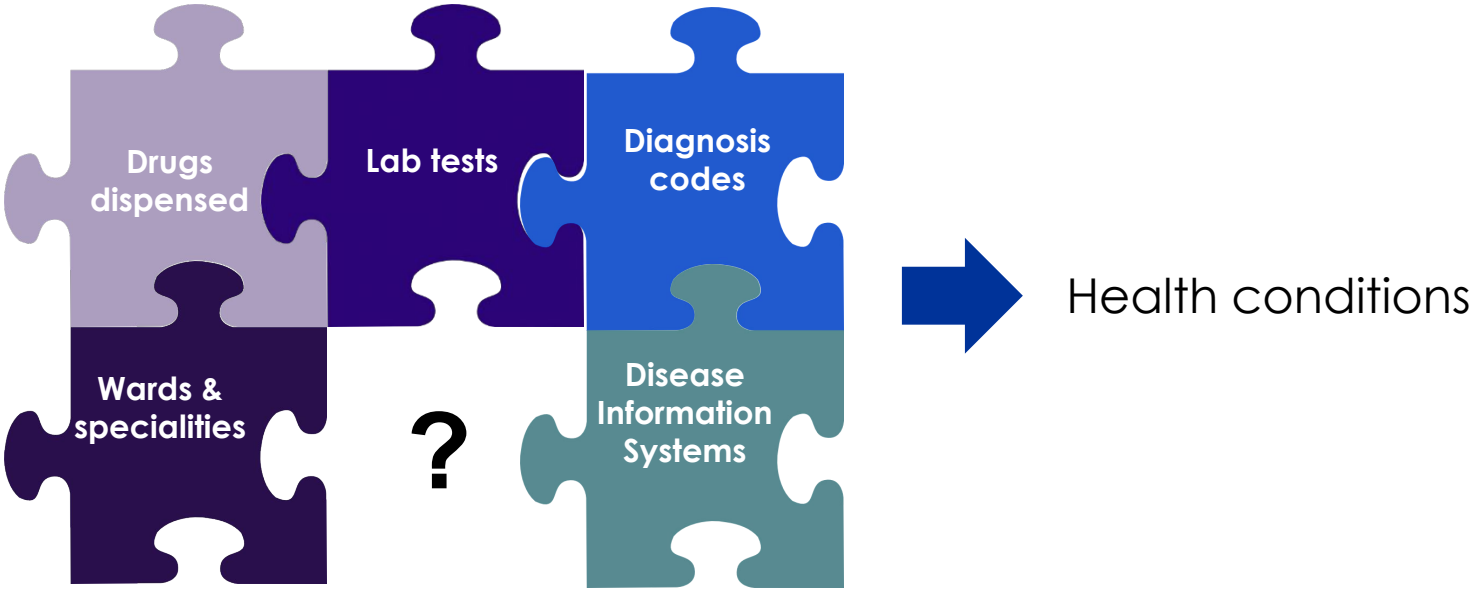
Total Births (including those with unknown mothers): **77078**

References: 3

Episodes

Episodes

- An **episode** is some kind of health condition that is inferred from the data, e.g. HIV, TB, pregnancy, etc.
- Made by SQL stored procedure
- One row per patient per episode
- Start date, end date, evidences



References: 3

Example: pregnancy

Example evidence list

Evidence	Source	Confidence level	Date
Rhesus test performed	Labs	High confidence	2015-01-23
Iron & Folate dispensed	Drugs	Supporting	2015-04-03
Has live birth record	Birth register	Outcome	2015-06-16
Has diagnosis code indicating live birth	Diagnosis codes	Weak-Moderate	2015-06-16

Rolled up into a single record per pregnancy

Patient	Pregnancy number	Start Date	End Date	Last Contact date	Evidence list	End date evidence list	Facility	Confidence
xxx	1	2015-01-23	2015-06-16	2015-06-16	Birth Record, Diagnosis code, Rhesus Test, Iron & Folate	Birth Record, Diagnosis code	MMH	0.95 ➔ High confidence

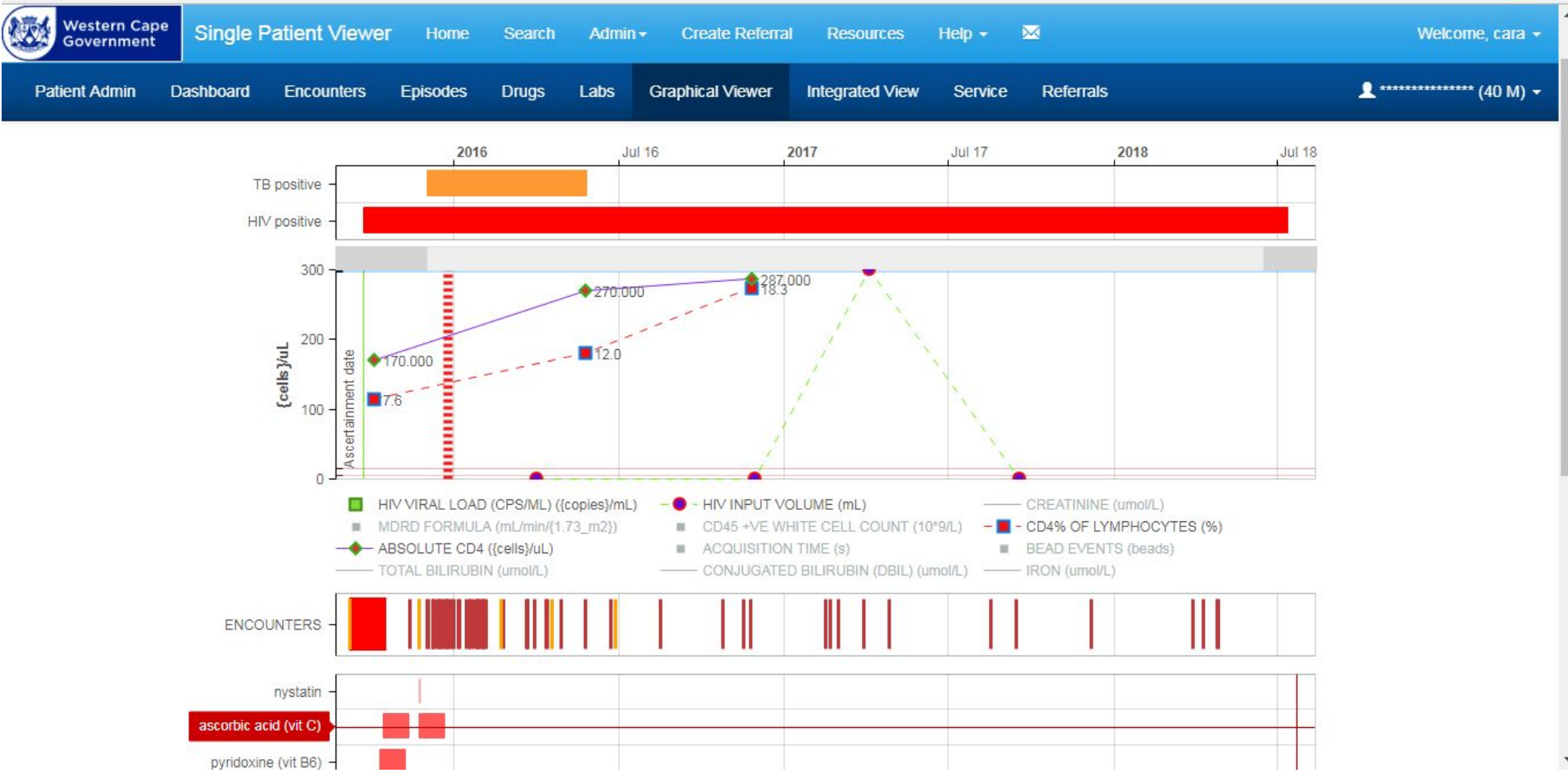
Single Patient Viewer

Single Patient Viewer

The SPV is the web app built by Jembi that allows doctors to:

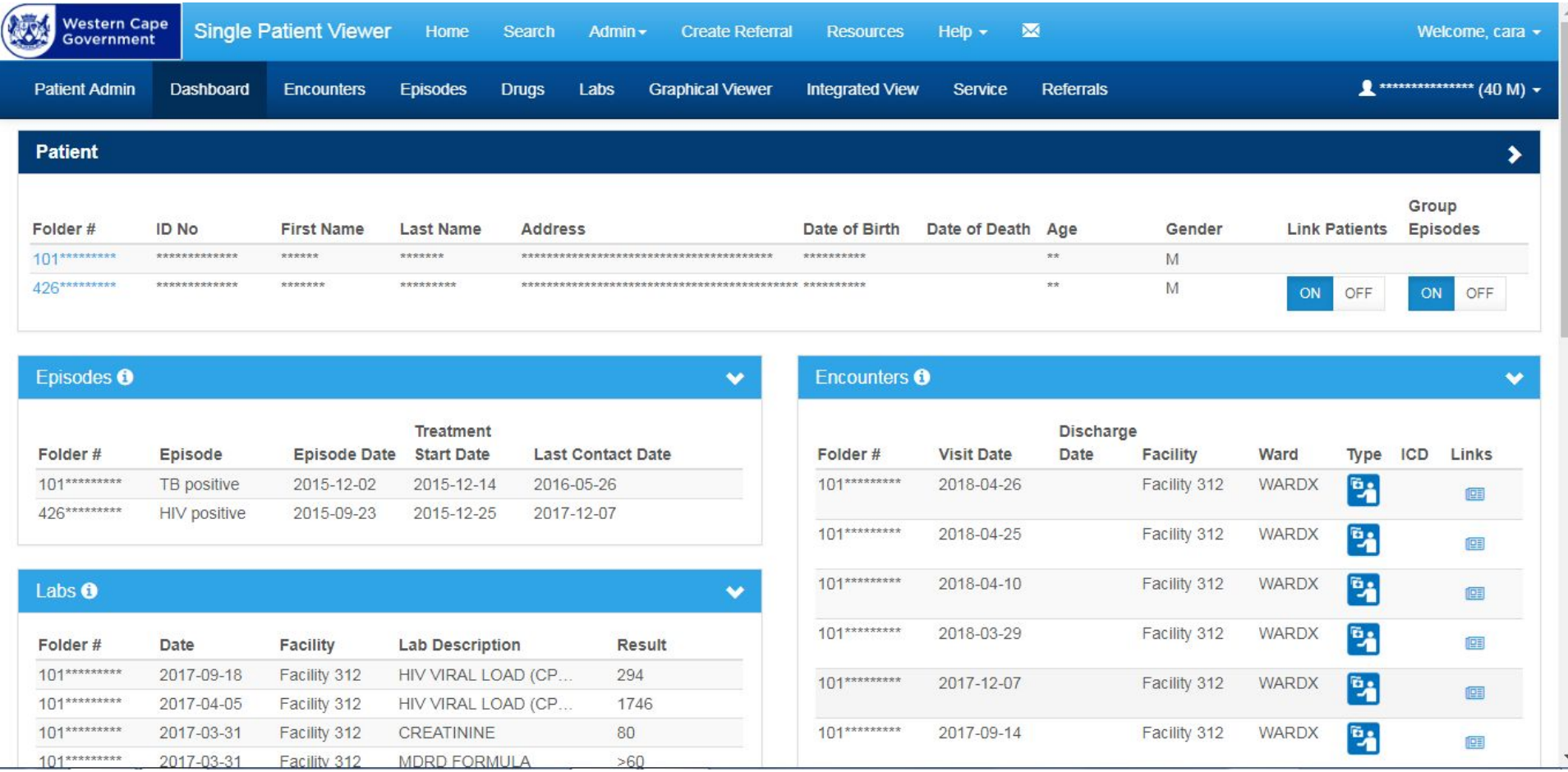
- Look up a patient using their name or folder number
- See their history, including lab tests, drugs, hospital admissions, diagnoses and visits in one place.

References: 3



SPV Graphical Viewer

References: 3



Screenshot of SPV patient dashboard

References: 3

Other uses for the Western Cape Data Centre

Cascade reports

- A list summarising patients at a facility with a specific health condition and their status
- Easy to see from the list who needs to be followed up (e.g. high viral load, abnormal pap smear)

References: 3

Monitoring and evaluation

- Having data allows us to measure how well our health system is performing
- We recently estimated from our data that almost a third of women diagnosed with cervical cancer in 2015 had no evidence of pap smears
- This is something the DoH can investigate

Alerts

Current:

- HIV patients on ARVS not responding to treatment.
- Alert NHLS so they can test for genotypic resistance.

Possible:

- Find a measles patient in our daily load.
- Automatically email every public sector doctor in the province to alert them to an outbreak.
- Handled by NICD.

Research

Anonymised data extracts for

- Academics
- Parliamentary committees
- Other organisations

Ethics and data governance

All too often, people view personal data as just another piece of information that they are working with — a commodity. However, this data represents a real person and should therefore be respected.

— Elizabeth Maxwell

Ethics and data governance

Ethical principles when dealing with patient data:

- Confidentiality
- Must benefit the patient
- Avoid any potential harms
- Do not exploit vulnerable populations (e.g. people in very poor health or living in poverty)

References: 3

Ethics and data governance

Laws:

- Protection of Personal Information Act (POPI)
  - Data must be stored and shared legally
- Promotion of Access to Information Act
  - Records must be kept of how all data is accessed and used

References: 3

Ethics and data governance

To ensure security:

- Database and table-specific access permissions
- Personal information only stored in one database and kept under strict permissions
- Firewalls
- Any data transferred is password protected
- Anonymisation and aggregation wherever possible

References: 3

Summary

- We do a different kind of data science in a special context
- Accountability, transparency and consequences
- Come from a history of division, misinformation, separation
- Every day, we do our bit to bring some unity

Thank you

References

1. The health and health system of South Africa: historical roots of current public health challenges by H Coovadia et al. 2009.
2. National eHealth Strategy, South Africa 2012 - 2017 by the South African NDoH. 2012
3. Provincial Health Data Centre Update to HIA Unit Meeting slides by Andrew Boulle et al, 2017
4. What Is Interoperability by HIMSS, 2013 https://www.himss.org/library/interoperability-standards/what-is-interoperability

Image sources

1. https://en.wikipedia.org/wiki/Bantustan#/media/File:Bantustans_in_South_Africa.svg
2. http://maps-africa.blogspot.com/2012/05/south-africa-map-pictures.html