An Overview of the Data Governance Framework at the PHDC

Nicki Tiffin, on behalf of the PHDC

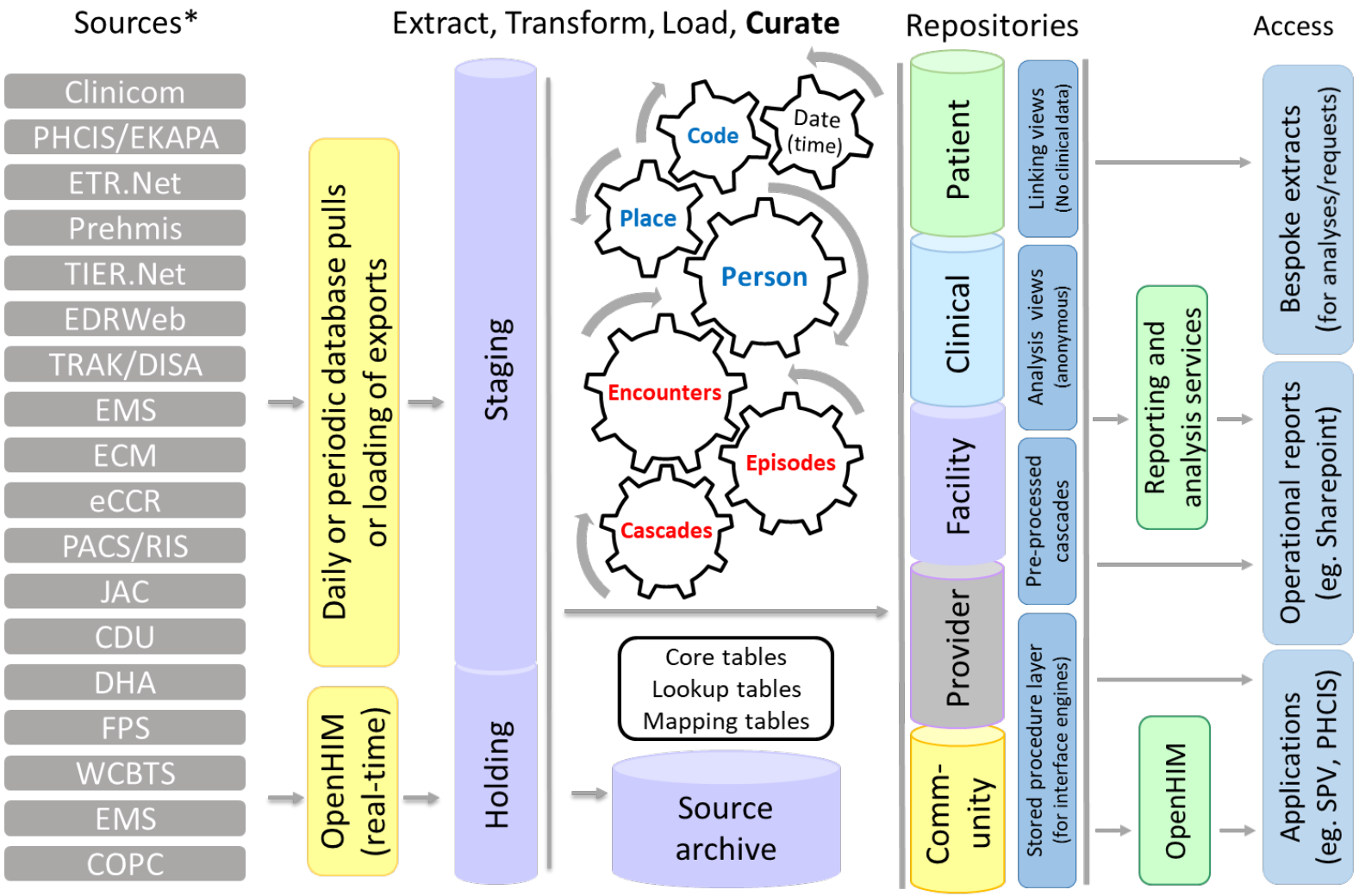
Western Cape Government
Health

Aims of the Provincial Health Data Centre

IMPROVED PATIENT OUTCOMES

- Continuity of Care
- Operational reporting
- Decision support system for health management
- Epidemiological and other research support

High level architecture



Considerations for data governance

- Data are collected from patients seeking health care.
- Health data are sensitive and personal data.
- Implicit agreement between patients and providers that data are provided for the **sole purpose of provision of health care**.
- The use of these data is governed by the **Health Act.**
- **No informed consent** is provided for any other use.

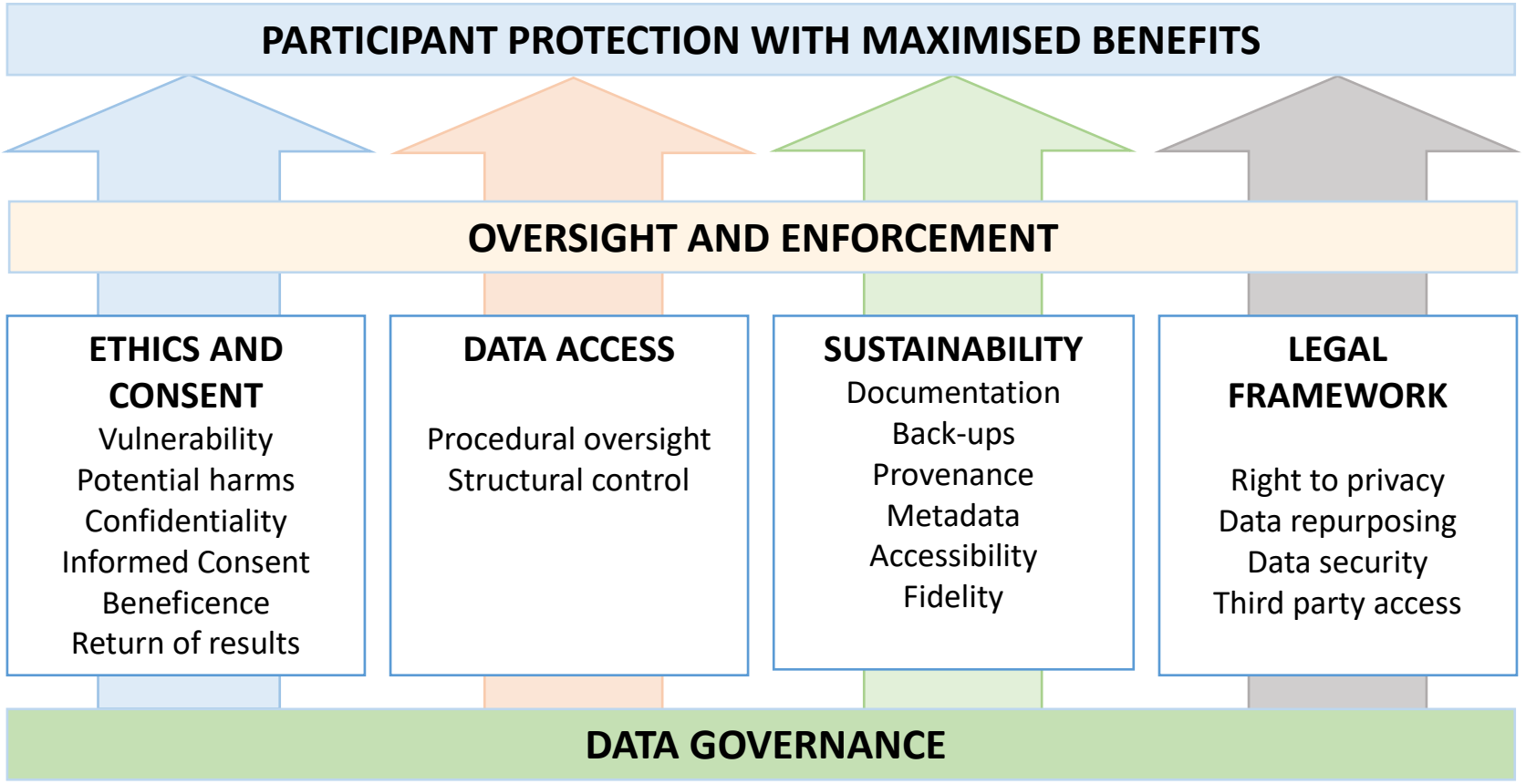
Consequences of data collation

- Increased data volumes
- Increased opportunity for evidence-based decision making
- Increased risk of data breach, privacy violation
- Vulnerable populations are disadvantaged in negotiating data protection
- These data *could* be useful for other virtuous purposes
    - evidence-based provision of care
    - understanding drivers of health
    - design of health interventions
    - M&E
    - academic research

Pillars of data governance at the PHDC

1. Participant Protection
   - Ethics, consent, confidentiality
2. Legislation
   - Health Act, POPI Act, PAIA
3. Data Access Controls
   - Procedural: Data access processes, checks and balances
   - Structural: firewalls, passwords, separate clinical and identifying information
4. Sustainability
   - Documentation, backups, provenance

Considerations for data governance



Tiffin N, et al. How to use relevant data for maximal benefit with minimal risk: digital health data governance to protect vulnerable populations in low-income and middle-income countries. *BMJ Global Health* 2019;**4:**e001395

Western Cape Government
Health

Pillars of data governance at the PHDC

1. Participant Protection

   Ethics, consent, confidentiality

2. Legislation

   Health Act, POPI Act, PAIA

3. Data Access Controls

   Procedural: Data access processes, checks and balances
   Structural: firewalls, passwords, separate clinical and identifying information

4. Sustainability

   Documentation, backups, provenance

Potential onward use of data:

- Provision of health care to individuals – primary use, clinical care
- Provision of better health care at a population level – operational use by the Dept of Health
  - Where do we need to direct resources? BoD
  - How can we link patients to appropriate care? Reports, line-listings
  - Who is "slipping through the cracks"?
  - What interventions are required – epidemiological analyses
- Operational use by NGOs, service providers providing health care in collaboration with/on secondment to the Dept of Health
- Research use by academics

Legislation, legal compliance

POPI Protection of Personal Information act

Health data are **Special Personal Information**
Responsible Party: Department of Health
Primary purpose of data collection: Provision of health care

Secondary data use/data repurposing/ data sharing not allowed (unless with informed consent)

HEALTH CARE ACT Patient confidentiality

PAIA Promotion of access to information act
Record keeping of how individuals' data are used
- records of all data access

Application of the POPI Act

The **Protection of Personal Information Act (POPI),** South Africa (equivalent to GDPR in EU)

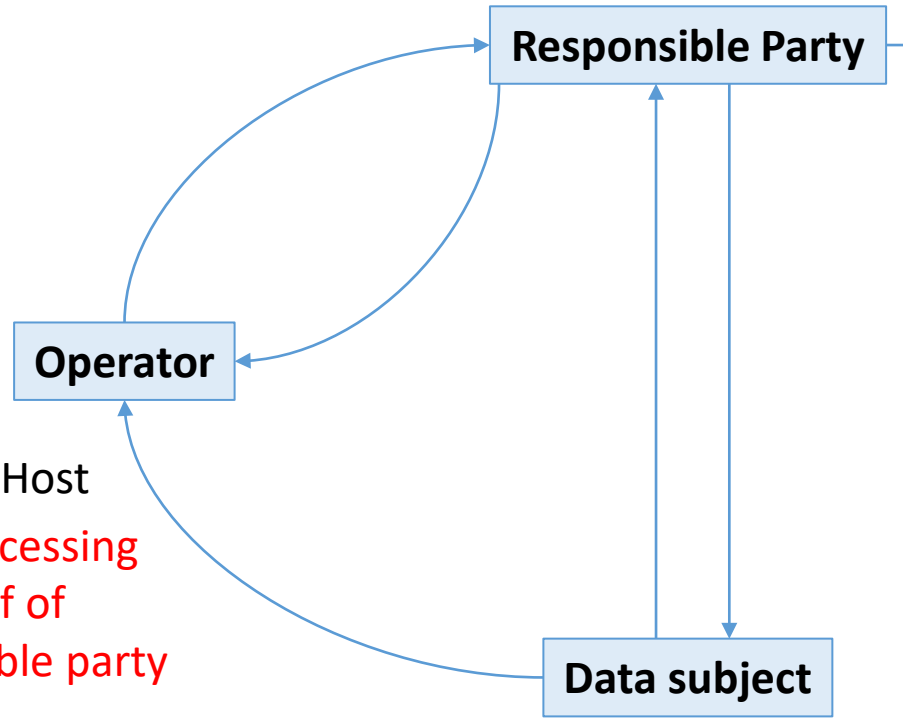
- Governs the use of personal information
- Synergises with the Health Act, upholds patient confidentiality

- Identifies “special” data, which include health data
- Governs international transfer of data

- Distinguishes identified data vs anonymised data
- Recognises informed consent

ROLES

e.g Government health dept
Academic researcher
Mobile phone company
Bank



e.g Researcher
Marketing company

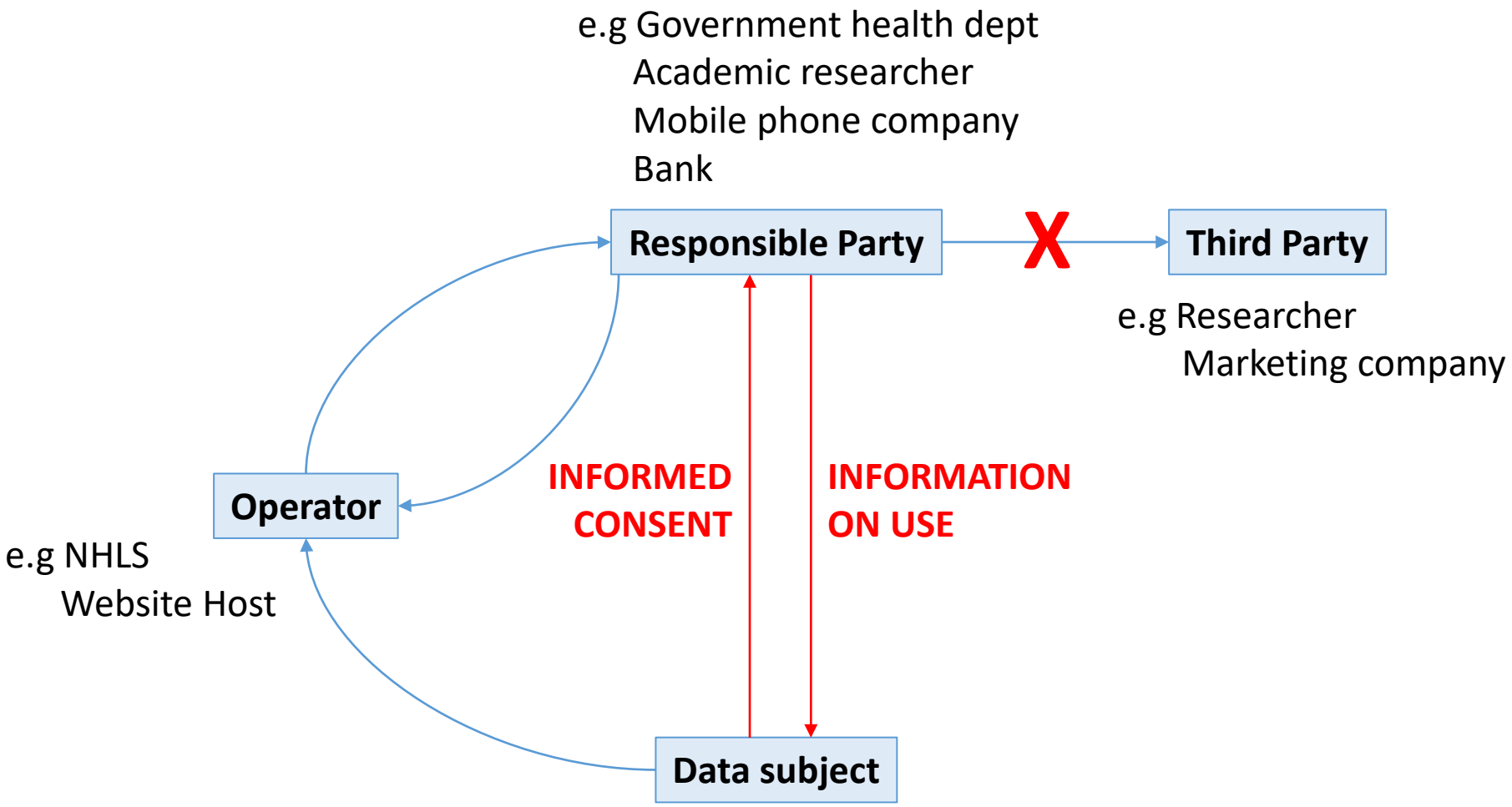
e.g NHLS
Website Host
Data processing on behalf of responsible party

"PROCESSING"
Collecting, transferring, storing, analysing

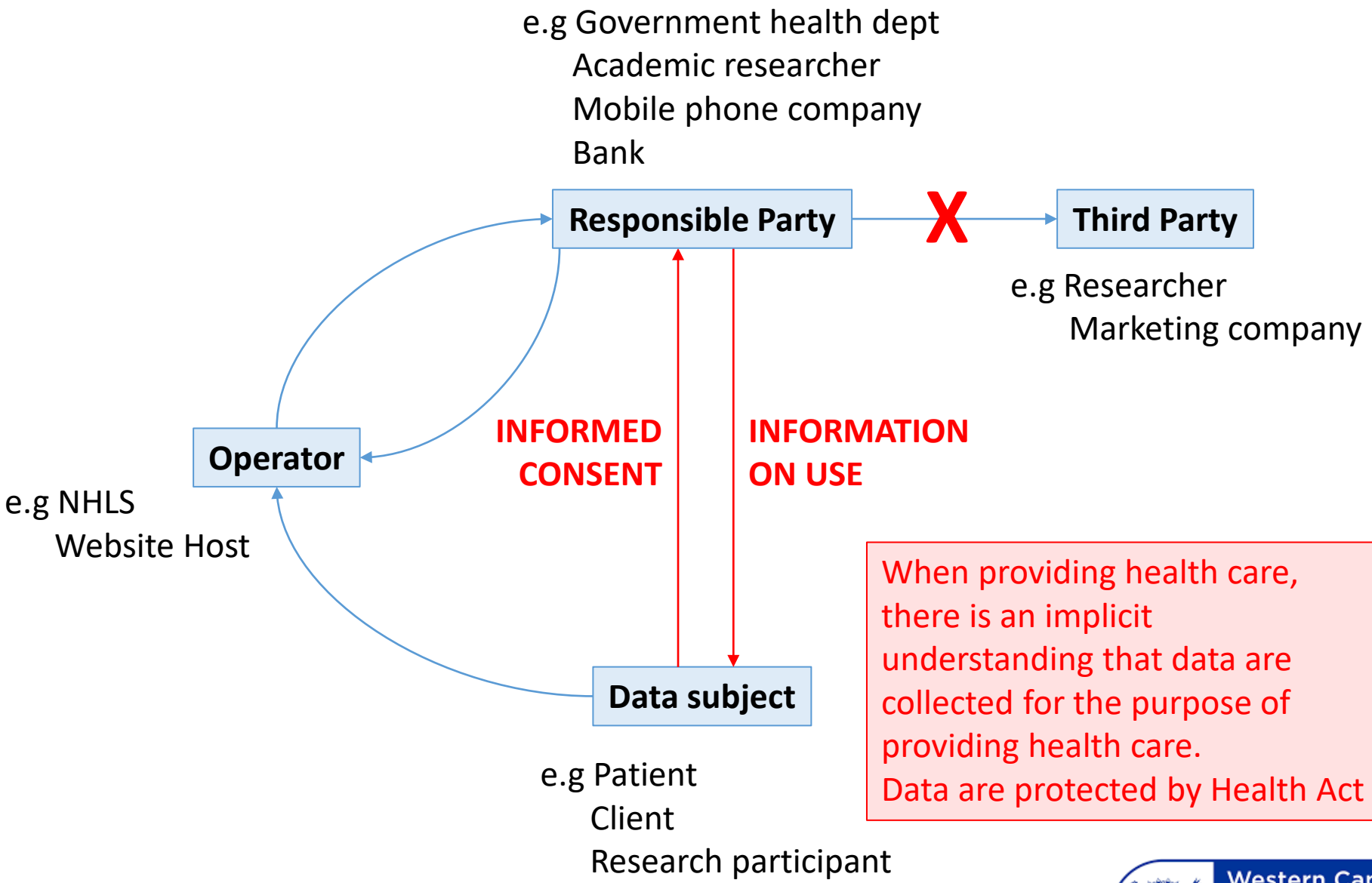
e.g Patient
Client
Research participant

Western Cape Government
Health

IDENTIFIED DATA



IDENTIFIED DATA



When providing health care, there is an implicit understanding that data are collected for the purpose of providing health care.
Data are protected by Health Act

What is allowed under POPI Act and Health Act?

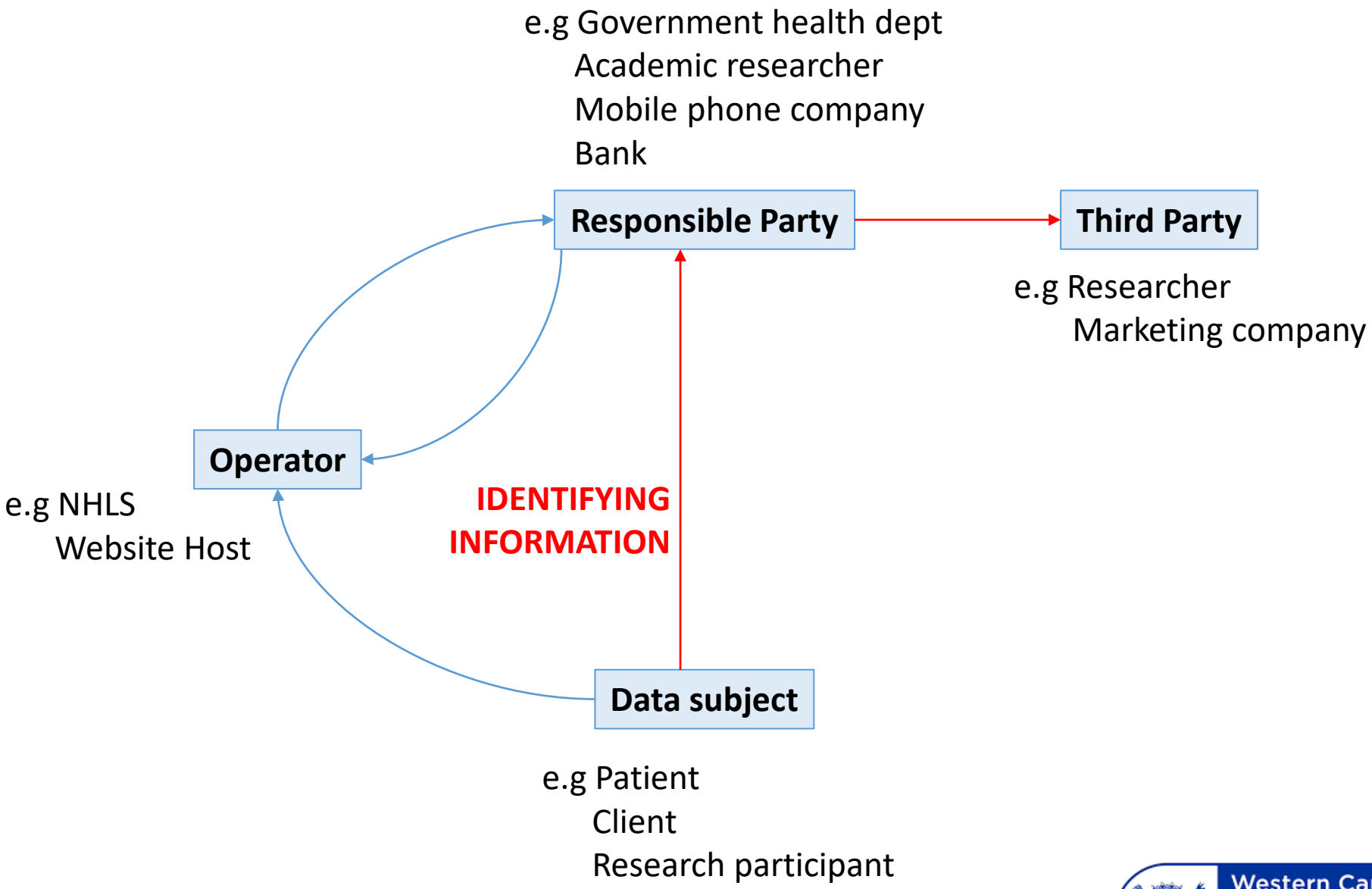
- Provision of health care to individuals – primary use, clinical care

- Provision of better health care at a population level – operational use by the Dept of Health
  - Where do we need to direct resources? BoD
  - How can we link patients to appropriate care? Reports, line-listings
  - Who is "slipping through the cracks" – types of patients?
  - What interventions are required – epidemiological analyses

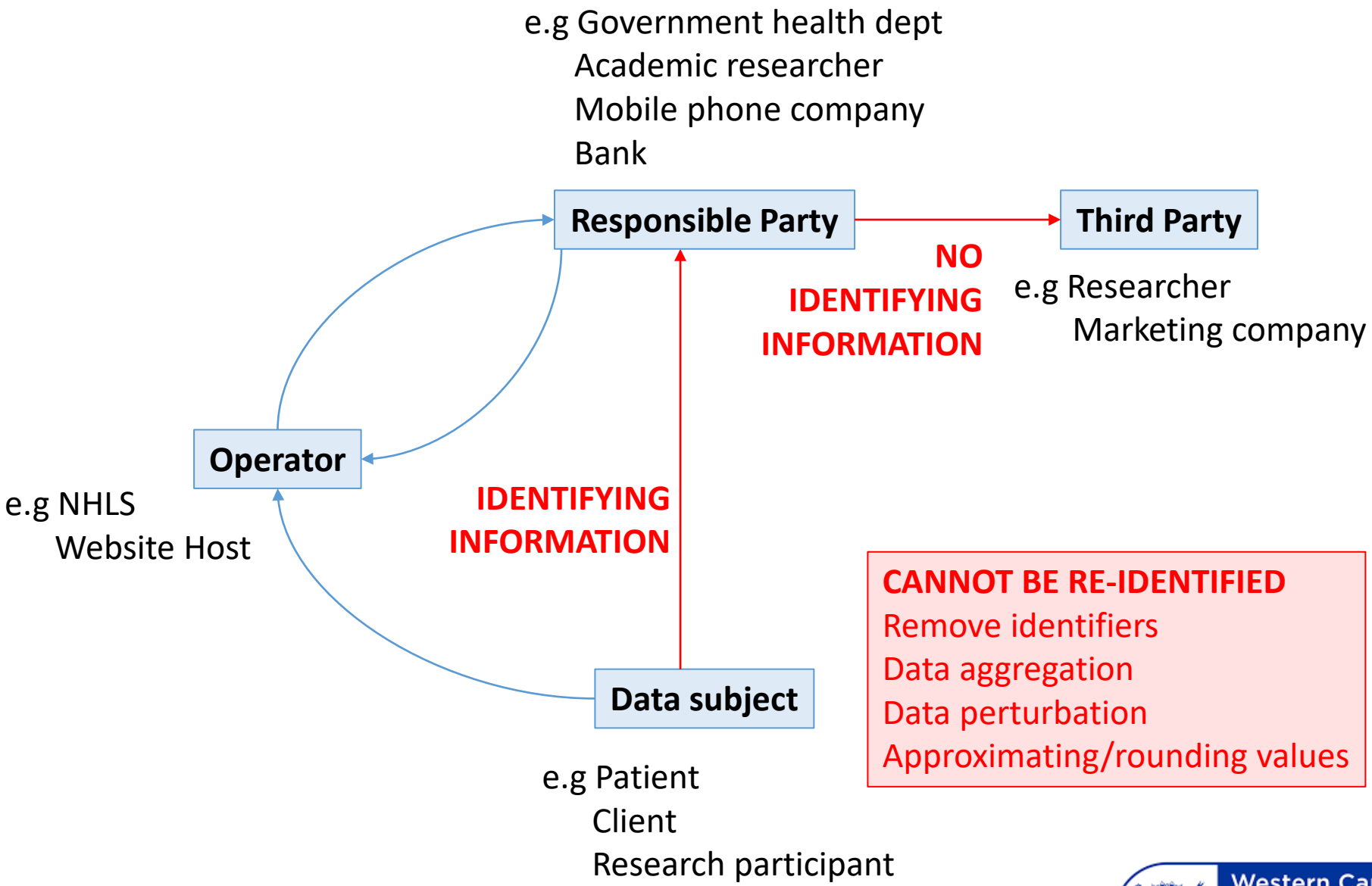
- Operational use by NGOs, service providers providing health care in collaboration with/on secondment to the Dept of Health, acting as Operators on behalf of Dept of Health (MOUs in place)

- Research use by academics

ANONYMOUS DATA



ANONYMOUS DATA



Can data be re-identified?

Data perturbation

- Hide, round off or alter dates eg. Year of birth; Age scale e.g. days for neonates, weeks for newborns, months up to 2yrs, years thereafter)

- 'time to' events in days from index event.
  eg. Time in days to death after admission

- Round off numbers (e.g. birthweights)

Trusted third party stewardship

Data aggregation

Geographic regions, shapes, suburbs. **No dots on maps**

Can data be re-identified?

e.g.
A 37-year old lady with epilepsy attending a particular clinic on given dates.

A 12 year old, female, grade 6 learner who lives at 1 Green Street, Townsville.

Do not manage de-identified data in the same way as truly anonymised data, because they can be re-identified.

What is allowed under POPI Act and Health Act?

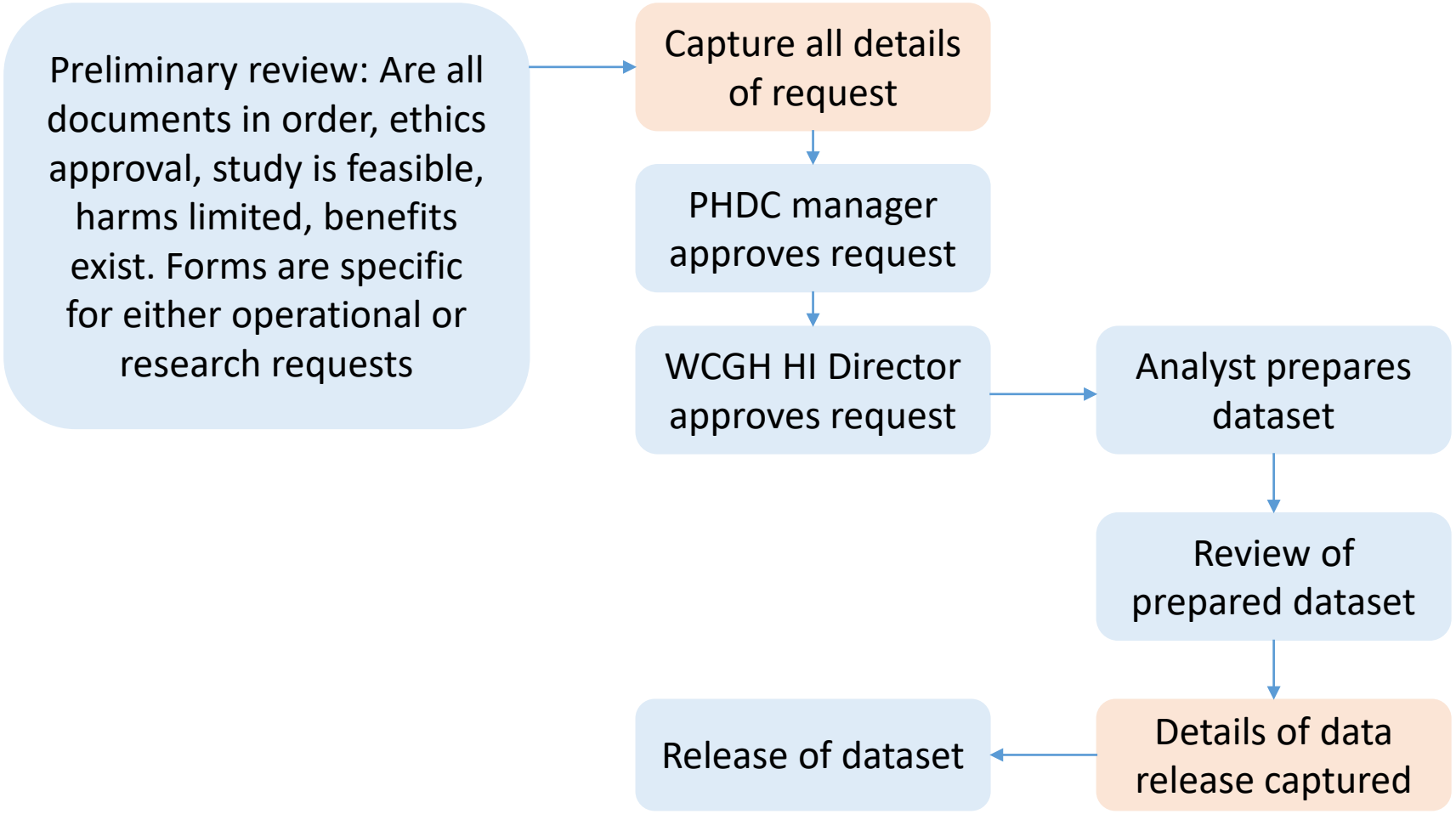
- Provision of health care to individuals – primary use, clinical care

- Provision of better health care at a population level – operational use by the Dept of Health
  - Where do we need to direct resources? BoD
  - How can we link patients to appropriate care? Reports, line-listings
  - Who is "slipping through the cracks" – types of patients?
  - What interventions are required – epidemiological analyses

- Operational use by NGOs, service providers providing health care in collaboration with/on secondment to the Dept of Health, acting as Operators on behalf of Dept of Health (MOUs in place)

- Research use by academics

Oversight for data access requests



Western Cape Government
Health

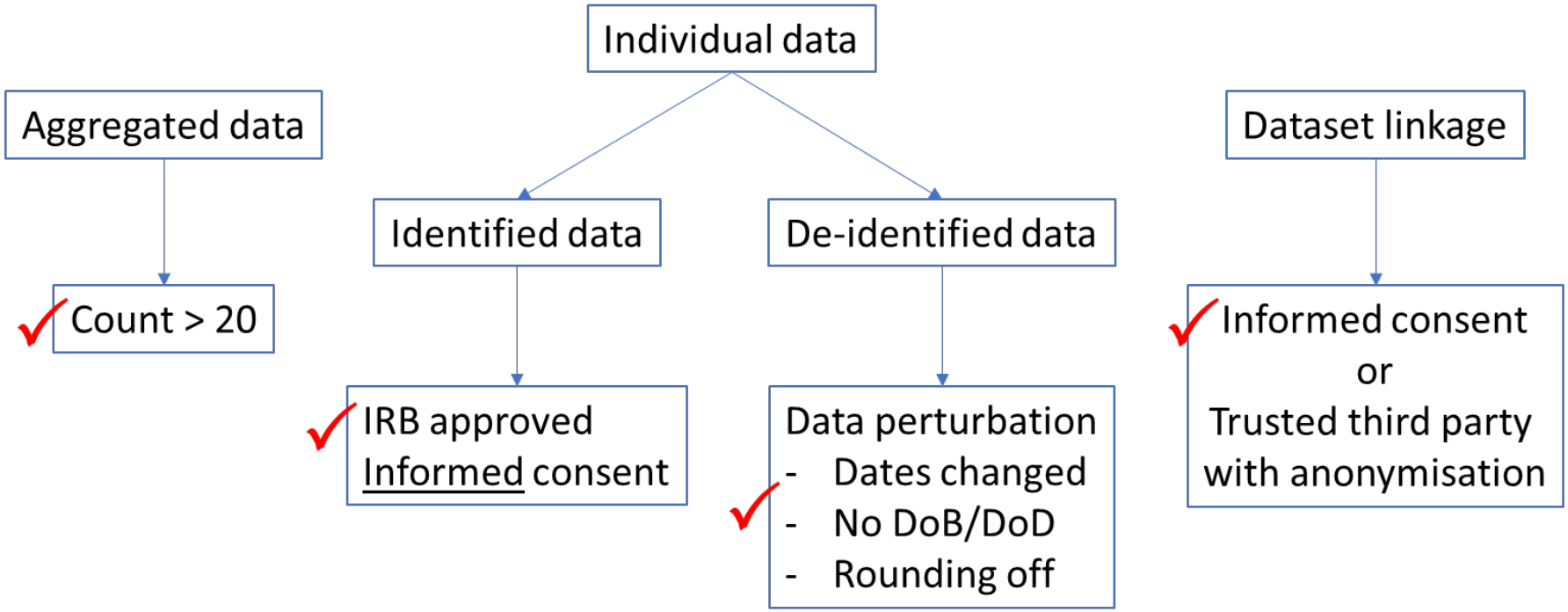
Oversight for data access requests

Review of external research requests for data:

- Does the study have ethics approval from a recognised board?
- Does the research protocol adequately describe accessing data from the PCDC/WCGH?
- Is there informed consent for identified data?
  - Read the information provided to participants, does it describe accessing electronic data from WCGH/PHDC?
  - Read the consent form template, do the participants consent to the access of their electronic health data from WCGH/PHDC?
- Is the study feasible?
- Do we have the data?
- Will the findings be accurate (the data are inherently biased)?
- Are anonymised data able to be re-identified?
- Could data result in exposing individuals or communities to harms (e.g. ‘dots on houses’, neighbourhood stigma)
- Are the data requestors somehow involved in the provision of health care in the Western Cape and understand the health systems and patients behind the data?

Data privacy ↔ Data use to improve health

EXTERNAL RESEARCH REQUESTS



Western Cape Government
Health

Some specific examples

It's often possible to find solutions to facilitate appropriate data use without violating patient privacy or legislation

- **Consent was given by participants for "access to information in medical folder" for research before there was a PHDC.**

Concern: Folders only provide info from attendance at a specific facility, whereas PHDC data are from all facilities, and many different types of data

Approached Provincial Ethics Review Committee for advice

Solution: Ask researcher to submit an amendment for review by their Ethics Review Board, for approval of the wider scope of data

Some specific examples

- **Researchers from different sectors (e.g. other provincial departments) want to link health data to their own/other datasets**

Concern: Under POPI Act, identified data collected for other purposes cannot be shared to third parties without consent, but identifiers are needed to link datasets

Solution: Use a trusted third party not involved with either dataset, nor invested in the study, to receive both datasets, join them and then de-identify and perturb the final dataset. Original datasets are then destroyed by the trusted third party.

A detailed MOU is prepared and signed by all parties, describes exactly how linkage will work and includes commitment to confidentiality and data destruction by the trusted third party

Some specific examples

- **Clinical researchers want to do research on data from their own patients, linked to additional collected data**

Concern: Without consent, identified data cannot legally be provided

Solution: Researcher provides their own bespoke dataset with identifiers to the PHDC;
then the PHDC links to additional health data, anonymises and perturbs the dataset sufficiently so patients can't be re-identified using the researcher's data they are holding

Some specific examples

- **Clinical researchers want to do a folder review of their patients to validate their health status as defined by PHDC data (e.g. episodes assigned). Then they want to do research with the PHDC data for their patients.**

1. Under an **operational data request**, the PHDC provides identified data to the clinician to review and validate PHDC data. The clinician only sees data for patients already in their care. They return the curated information to PHDC and don't use it further. This work helps the PHDC improve its data resources and improve its outputs.

2. The clinician separately requests anonymised research data for the epidemiological analysis, with ethics approval in place – **research data request**.

Some specific examples

- **WCGH partners may provide clinical care on behalf of WCGH, governed by formal agreements, and may want to also do research without consent using data they collect or that are provided by PHDC for the purposes of providing clinical care.**

Concern: Data provided by patients or to partners for the purpose of providing patients clinical care cannot be repurposed for research purposes without explicit consent.

Often the partner is acting for WCGH, so the patient believes they are providing their data in order to receive health care from WCGH

Some specific examples

Solution: The provision and receipt of data must be clearly identified to be operational (provision of care ONLY), or research.

Operational data exchanges may include patient identifiers as needed to provide care.

Separately, the WCGH partner can apply for research access to PHDC for an anonymised, perturbed dataset representing the data from these patients, with approval from an ethics review board.

Thanks to:

The Provincial Health Data Centre Team lead by Andrew Boulle

Melvin Moodley (and previously Tony Hawkridge) at the Health Impact Assessment Directorate at WCGH

The National Institute of Child Health and Development (NIH, USA): B-Positive R01HD080465
The Bill and Melinda Gates foundation: The African Health Information Exchange: OPP1164272
National Human Genome Research Institute (NHGRI), National Institutes Of Health (OD) H3ABioNet award, number U24HG00694
Wellcome CIDRI-AFRICA grant (203135/Z/16/Z)

Western Cape Government
Health