Engagement Roadshow on the NMRC Research Data Governance and Sharing Framework

20th and 21st September 2023

Framework for Research Data Governance and Sharing

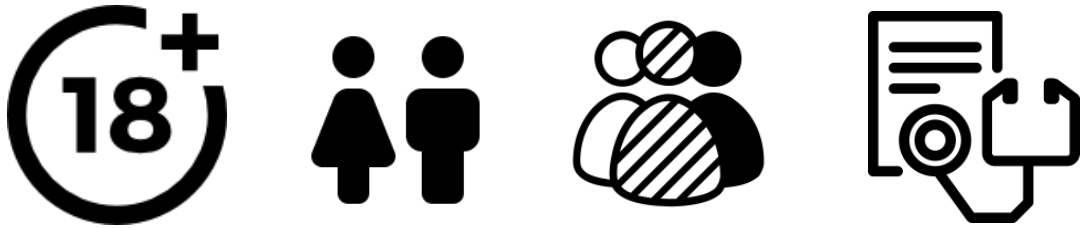
- NMRC developed a Framework for Research Data Governance and Sharing, which was endorsed by the HBMS EXCO in October 2014.
- Feedback was gathered from the institutions and supplemental baseline requirements were further developed.
- The framework requires publicly-funded project to:
  a) Allow open access to peer-reviewed publication. This is applicable to all ongoing and new projects from May 2015 onwards.
  b) Share final research data for projects seeking funding of S$250,000 direct cost and above.

Research Data Sharing

- Promotes open scientific inquiry that expedites the translation of research findings into knowledge and applications, to improve human health and healthcare.
- Promotes novel research that may not have been envisioned by the original researchers, allows synergies from multiple data sources, and enables the sharing of unique data that cannot be easily replicated (such as studies on unique populations).
- Allows other researchers to leverage on past research findings and facilitates collaboration to make research more robust.

Research Data to be Shared

- Researchers are required to share the final research data and its associated metadata. This is applicable to all type of studies.
- With respect to clinical data, if there are reasonable limitations to sharing, a minimum set of clinical data, including basic demographics (namely, age, gender and race), and disease or condition (diagnosis), should be shared, unless the information is not collected in the research.



Age **Gender** **Race** **Diagnosis**

Final Research Data

- Final Research Data refers to recorded factual material commonly accepted in the scientific community as necessary to document and support research findings.
- The general guiding principle to what constitute final research data would be what researchers themselves find sufficient to validate research findings and support future research work.
- Data standards and formats are not stipulated at this stage.
- Sharing of partial data is possible, if restrictions are deemed reasonable.
  - Such limitations (e.g. omission of certain data, datasets only comprise data from 60 out of 80 subjects, or only anonymized data is collected) should be made known in the metadata so that data requestors/public are aware.

Data Owners' Responsibilities and Rights

- Data contributors/data contributing institutions, who are the data owners, will be responsible for the integrity, quality and accuracy of the final research data submitted.
- Data owners continue to have the rights to use their research data for future research and for sharing based on their discretion.

Timeliness in Data Sharing

- Data is to be made available at the earliest feasible opportunity.
- Delays or restrictions on data sharing may be justified to gain intellectual property protection or to further develop technology for public benefit.
- However, such limits should be minimised as far as is feasible and be indicated in the Data Sharing Plan.
- Data should be made available at the earliest feasible opportunity when the restriction no longer applies.

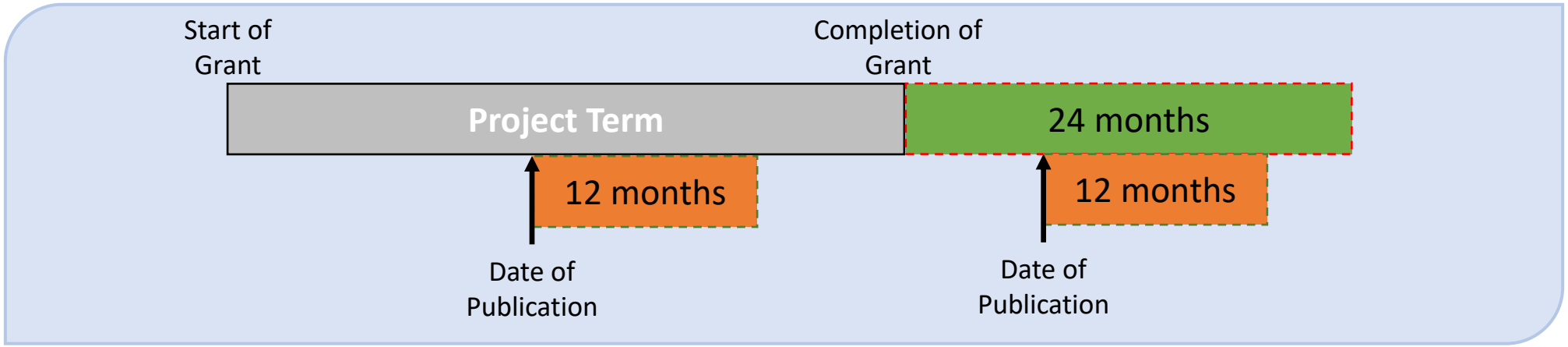
Timeframe for Sharing Research Data

Dataset used in publication

No later than 12 months after official date of publication

Dataset not used in publication

No later than 24 months after grant completion



Extension of deadline may be sought if there are strong justifications for the extension (e.g. manuscript under review, intellectual property protection) and evidence of the same, subject to NMRC's approval.

Human Subjects and Privacy

- It is the responsibility of the investigators, their Institutional Review Board (IRB), and their institutions to protect the rights of study subjects and the confidentiality of the data, and to comply with applicable legislation such as the Personal Data Protection Act and the Human Biomedical Research Act, to name a few.
- Data shared must be shared in a robustly anonymised form, i.e. free of identifiers that would link to, or lead to deductive disclosure of the identity of the human subjects. The anonymisation standards used shall be equivalent to or more robust than the government recognised anonymisation standards.

Informed Consent for Data Sharing

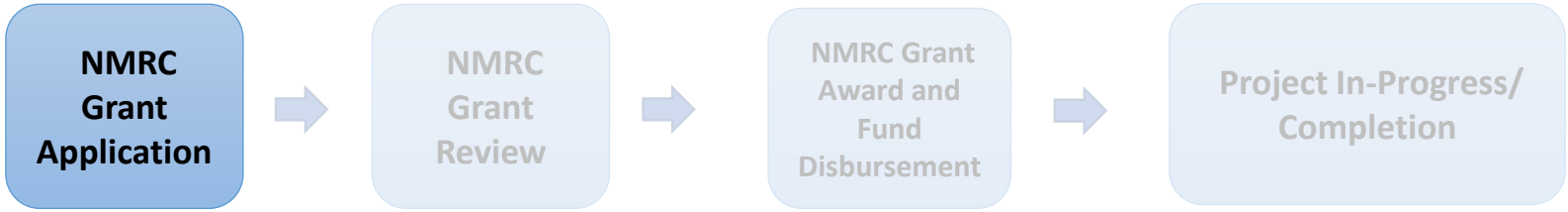
- Prospective participants should understand what is expected to happen with their data, and the data shared (if any) will be in anonymised form.
- Informed consent to include, where appropriate,
  - explicit provision for sharing of data for wider or future research use,
  - subject's approval to be contacted should the research or future research show that clinical intervention is required,
  - consent to collect data from national-level databases (e.g. clinical) in anonymised form.
- Informed consent should be sought with option to opt out.
- For past studies where informed consent for sharing of data for wider/future research use was not included, this could be stated in the data sharing plan for consideration of data sharing to be waived.
- Linkage consent would facilitate linkages to other datasets in the future, but requestors may seek IRB approval for waiver of consent downstream.
- Investigators and their institutions may request removal of data on individual participants from the repository in the event that consent is withdrawn. However, data that have already been distributed to approved users for research will not be able to be retrieved.

Storage of Final Research Data

- A copy of the Final Research Data shall be stored in the NMRC Research Data Repository after they have been anonymised.
- Final Research Data shall be stored in the Research Data Repository for a minimum duration of ten (10) years from each Deadline, and shall be subject to the Research Data Repository's archival protocol.

Implementation of the NMRC Research Data Governance and Sharing Framework

Plan and Budget for Cost Related to Managing Research Data



At the grant application stage, PIs are required to plan and budget for cost related to managing research data.

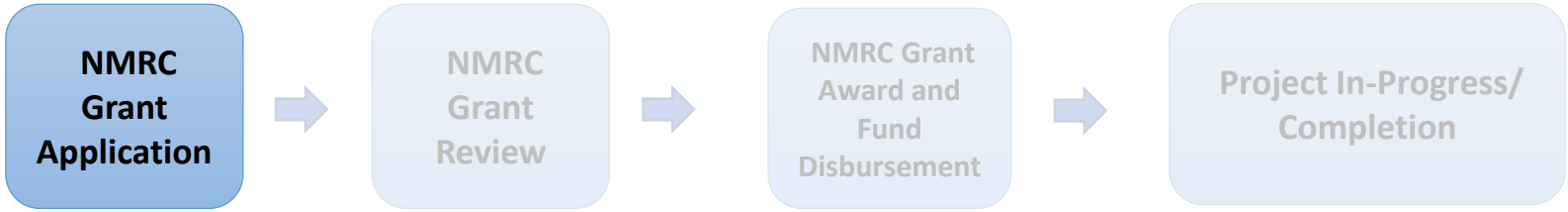
Type of cost to consider that may be included in the budget within the NMRC grant application:

- Cleaning, formatting, anonymising the final research data generated from the proposed research

Note

1) No cost will be incurred by PI/HI to deposit the final research data from the funded research in the NMRC Research Data Repository.
2) Institution core support manpower (if any) may potentially be supported by indirect cost or other institutional funding

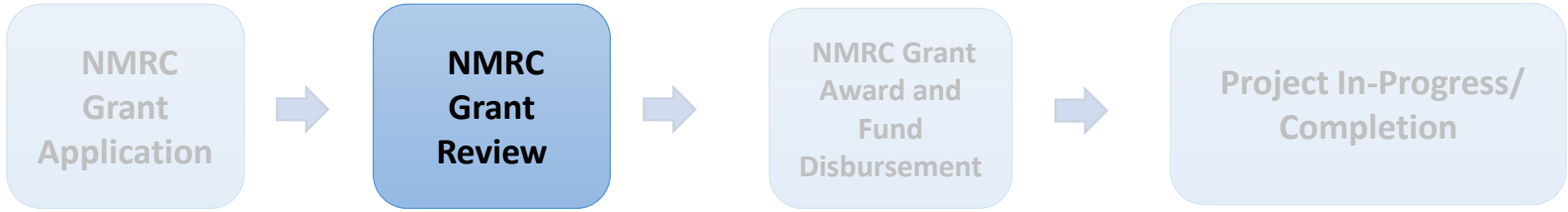
Submission of a Data Sharing Plan as part of the Grant Application



PIs are required to complete the **Data Sharing Plan template** as part of their grant applications, consisting of 4 sections:

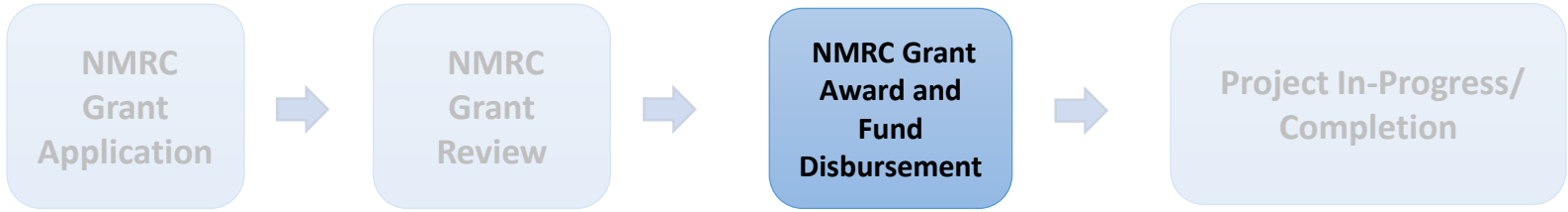
- **Description of Research and Research Data**
- **Restriction on Data Sharing**
  - ✓ *Restriction to data sharing may be due to participant confidentiality, consent agreements or Intellectual Property Rights.*
- **Data Use Limitation**
  - ✓ *Data use limitation should be minimised as far as possible. The limitations should be based on the terms of informed consent of the study participants from whom the data are generated, or based on valid justifications.*
- **Institutional Review Board (IRB) Endorsement (For a Research Study that involves human subjects)**

Submission of a Data Sharing Plan as part of the Grant Application



- The completed data sharing plan will be reviewed by the NMRC's review panels or equivalent body, and the IRBs (for studies involving human subjects).
- Reviewers are required to assess the reasonableness of the (i) data sharing limitation, and (ii) data use limitations (e.g. is it justifiable to limit sharing due to intellectual property rights protection, is it reasonable to limit data use to certain research areas)
- The Data Sharing Plan is assessed independently from the proposed research. However, a satisfactory Data Sharing Plan is required prior to the award of the grant.
- Upon the review and assessment, the NMRC's review panel or equivalent body may request the Data Sharing Plan be revised at the budget scrubbing stage. The revised Data Sharing Plan will be re-assessed by the reviewer/review panel if required.
- A **guidance document** will be provided to the reviewers who are required to review the data sharing plans as part of the grant applications.

Requirements prior to Project Commencement



If applicable, PI and HI are to obtain IRB endorsement on the data sharing plan in the ethics approval application for studies involving the use, collection and analysis of human subject data before the work requiring IRB approval may commence.

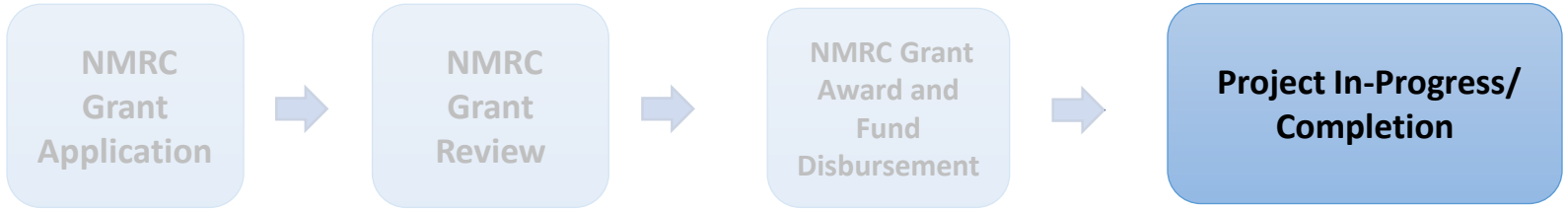
Review of Data Sharing Plan by IRB

Review by IRB

For studies involving the use, collection and analysis of human subject data, an Institutional Review Board (IRB) has to review and verify before work requiring IRB approval commences, that the data sharing plan in the ethics approval application is consistent with the informed consent of study participants from whom the data will be obtained, and there is adequate protection for the study participants involved in relation to data sharing.

- The IRB review should include assessment/consideration on
  - Consistency of data sharing plan with the informed consent procedures;
  - Adequacy of anonymisation standards used to protect human subject involved;
  - Sensitive data that may potentially lead to stigmatization;
  - Appropriate parameters for use of data through the use of data use limitations that could minimise the potential for harm.
- The IRB's decision will also take priority over the assessment of the NMRC Review Panel with regard to human subject protection.

Update on Data Sharing Plan and Deposit of Final Research Data



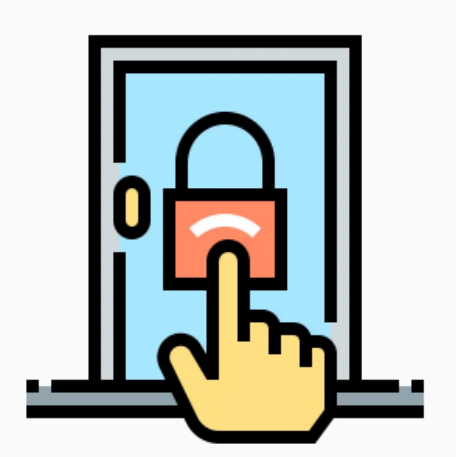
- PIs are to include an update on the data sharing plan (if there is any change) as part of the Final Report, that is currently submitted within three (3) months following the end of the project term.

- A copy of the anonymized final research data is to be **deposited at the NMRC Research Data Repository according to the stipulated timeline.**

Note:

- For projects that deal with datasets involving full NRIC/FIN, mapping tables (by respective data sources) should be retained and provided if there are approved data access requests that require data linkage.

Availability of Data on the Research Data Repository



The following will be openly available to serve as data catalogue and inform the community the data available:

- Project Information
- Metadata of the final research data

Final Research Data (anonymised form) in the Research Data Repository may be made available to a data requestor upon request, subject to approval by the data contributing institutions.

Submission of Data Access Request (DAR)

Qualified investigators[1] are to submit a **Data Access Request (DAR)** that includes:

a) Research Use Statement[2]

b) Terms of Use[3]

Note:

- *DAR must be endorsed by the requesting Investigator's institution.*
- *DAR is investigator-specific and with specific use case.*

[1]**Qualified Investigators** must be employees with a primary appointment in a Singapore public institution at a level equivalent to a Principal Investigator. Laboratory staff and trainees such as graduate students, and postdoctoral fellows are not permitted to submit data requests.

Submission of Data Access Request (DAR)

Qualified investigators[1] are to submit a **Data Access Request (DAR)** that includes:

a) **Research Use Statement[2]**

b) Terms of Use[3]

[2]**Research Use Statement** (cap at 2 pages) should include the following components:

- Objectives of the proposed research;
- Study design and analysis plan;
- Explanation of how the proposed research is consistent with the data use limitations of the requested dataset; and
- Funding source (including those awarded and those pending outcome).

Submission of Data Access Request (DAR)

Qualified investigators[1] are to submit a **Data Access Request (DAR)** that includes:

a) Research Use Statement[2]

b) **Terms of Use[3]**

[3]**Terms of Use** would include terms such as:

- To use the data only for the approved research;
- To protect the confidentiality of the data;
- To not attempt to identify or contact individual participants from whom the data were obtained;
- To not make copy or sell the data to any third-party entity or individual;
- To acknowledge in all oral or written presentations, disclosures, or publications the contributing investigator(s) who conducted the original study, the funding organization(s) that supported the work and the specific dataset(s);
- To agree to the listing of requestor's name and institutional affiliation, and the approved use of dataset.

Assessment of Data Requestor's Eligibility

NMRC will check the data requestor's eligibility before sending the requests to the relevant data contributing institution for assessment.

Review of Data Access Request

The DAR will be **reviewed by the Data Contributor's cluster/institution-based Data Access Committee/ Representative**.

Decisions to grant access should take into account whether the proposed use of the dataset is scientifically and ethically appropriate, and whether the request conforms to the NMRC Research Data Governance and Sharing Framework and is consistent with the Data Use Limitations prescribed for the dataset (including legal restrictions). For the avoidance of doubt, the above shall not in any way curtail the discretion of the data contributing institution's data access committee/representative or equivalent as to whether to grant access.

Note:

- Appeals for rejected request can be evaluated by the NMRC Data Access Oversight Committee (DAOC).

Approach for Collaborative Projects

- Data might be owned by different data contributors/institutions in a collaborative projects.
- The project team is encouraged to explore feasible arrangements to facilitate requests to access the research data.
- For example, a project-level data access committee comprising representatives of all data contributing parties could be set up for the lead PI's institution's data access committee to consult when granting approval to data requests, or the different parties could defer the decision on the study data to the lead PI's institution's data access committee.

Access to Requested Research Data

Data Contributors

Deposit a copy of the final research data in anonymised form

NMRC Research Data Repository

Note: Data will be retained for a minimum period of ten (10) years. Long-term/permanent archival plan will be worked out.

Requested data are transferred to TRUST/BRAIN for access upon data contributing institution's approval

The original data in NMRC Research Data Repository remains intact.

Note: If the approved data access request requires data linkage, data contributors will need to provide the mapping tables to facilitate the linkage.

Platforms for Data Access/Analytics

TRUST

For data request that requires linkages with other data

BRAIN

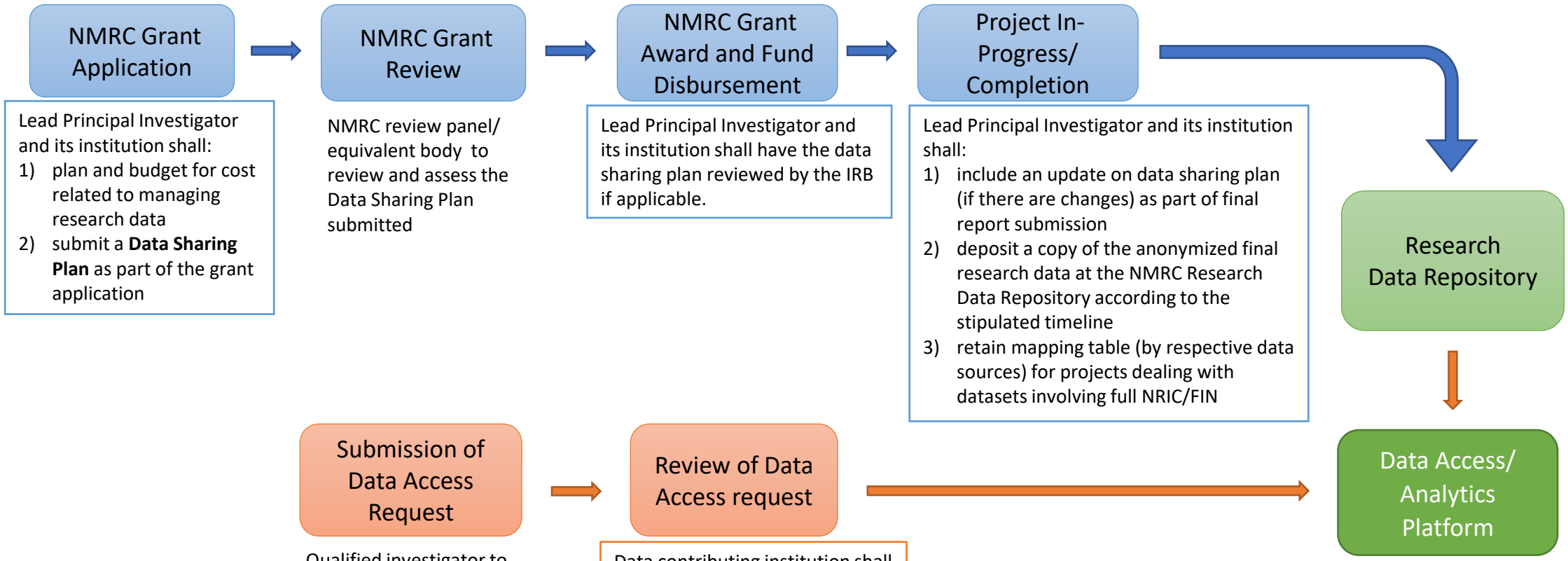
For data request that does not require linkages with other data

Note: Other platforms are not considered due to considerations such as data security.

Data Requestors

Access to requested data via BRAIN/TRUST in anonymised form

Impacts of the Framework Implementation on PI/HI



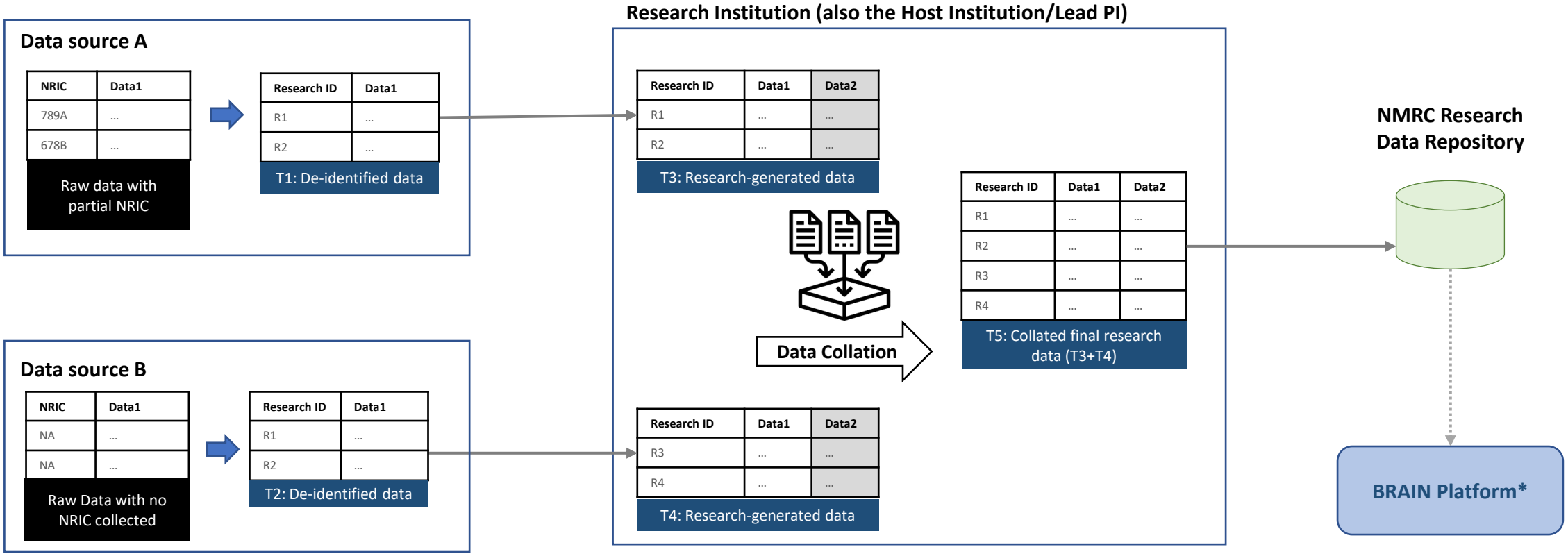
If approved, requested data are transferred from the research data repository to the data access/analytical platform(s) for access.

Requestors whose requests are rejected may appeal to the NMRC Data Access Oversight Committee.

Data contributor/ data contributing institution shall provide mapping tables to facilitate data linkage if the approved data access requests require data linkage.

Anonymization of Research Data

Anonymization Workflow for Projects that do not collect Full NRIC/FIN



Notes for Data Sources:

1) Data source will de-identify based on institution's current practice, and provide data with de-identified IDs (e.g. T1 with Research ID) to research institution for the conduct of research.
2) Mapping table is not required to be retained or provided as data fusion is not possible with projects that do not collect full NRIC/FIN.

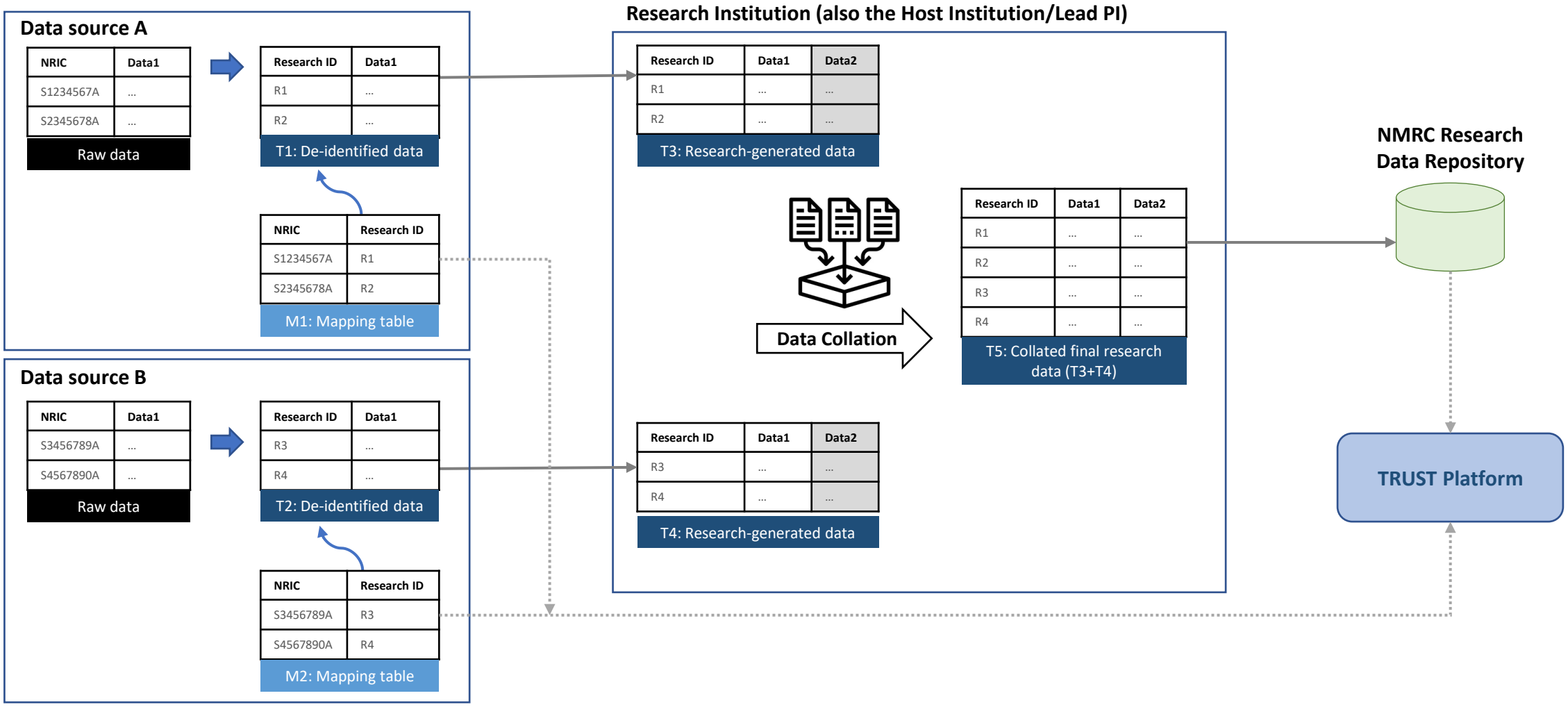
Notes for Research Institution:

1) Host institution/Lead PI is responsible to collate the final research data into a single, combined dataset.
2) All data must be anonymised, i.e. remove/suppress identifiable attributes of the data in accordance with government recognized annoymisation standards (or even more robust standards), prior to submission to the NMRC Research Data Repository (e.g. T5).

*** Data Privacy Measures**

1) Identifiers are not on board to the platform.
2) For data access (without data fusion as full NRIC/FIN is not collected), a use case specific hash is applied to the Research ID from the final research data before the anonymized data is made accessible via BRAIN.

Anonymization Workflow for Projects that Collect Full NRIC/FIN



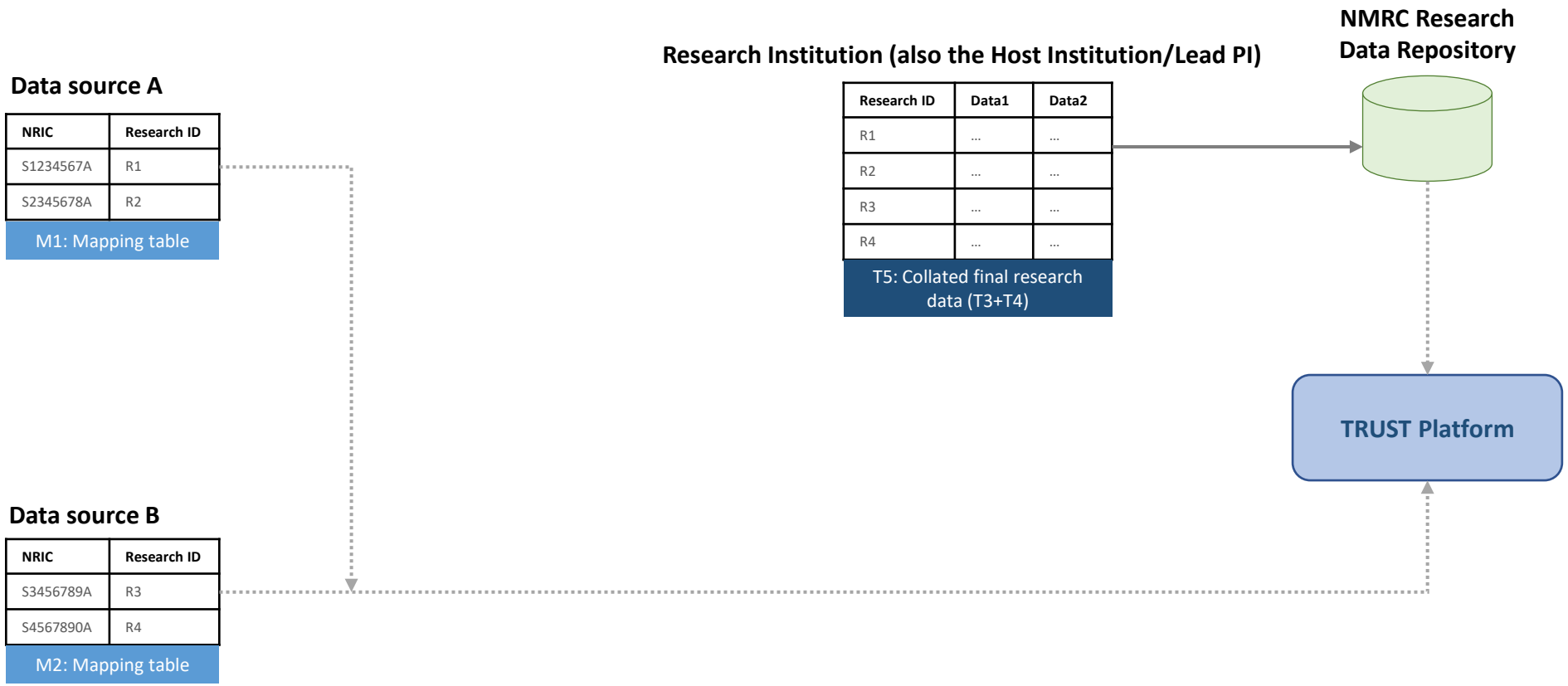
Notes for Data Sources:

1) Data source will de-identify based on institution's current practice, and provide data with de-identified IDs (e.g. T1 with Research ID) to research institution for the conduct of research.
2) Mapping tables (e.g. M1 and M2) are maintained by data sources, and are required to be provided if there are approved data access requests which require data linkage.
3) It is recommended to have a single party to maintain a collated mapping table for each project, if feasible.

Notes for Research Institution:

1) Host institution/Lead PI is responsible to collate the final research data into a single, combined dataset.
2) All data must be anonymised, i.e. remove/suppress identifiable attributes of the data in accordance with government recognized anonymisation standards (or even more robust standards), prior to submission to the NMRC Research Data Repository (e.g. T5).
3) The Research ID provided by data sources should be maintained and included as part of final research data to facilitate data fusion, in the event of such requests in future.

TRUST Data Fusion Workflow



TRUST Data Fusion Workflow

Data Contributor environment

MOH/TRUST environment

Research Institution (also the Host Institution/Lead PI)

Research ID	Data1	Data2
R1	…	…
R2	…	…
R3	…	…
R4	…	…

T5: Collated final research data (T3+T4)

NMRC Research Data Repository

First-level hash is applied to a copy of the final research datasets (Research ID) in the NMRC Research Data Repository before going into TRUST.
E.g., Research ID R1 -> TOK321

Data Privacy Measures

1) TRUST does not onboard identifier.
2) All de-identified IDs undergo two levels of hashing
- 1st : S1234567A -> ABC123
- 2nd: ABC123 -> XYZ0987
- XYZ0987 cannot be reversed to S1234567A

TRUST Platform

ABC123 De-identified & Cleaned Research Data

ABC123 De-identified & Cleaned Data from other data source

Second-level hash is applied (use case specific)

XYZ0987 De-identified & Cleaned Research Data

XYZ0987 De-identified & Cleaned Data from other data source

Hashed ID matches, thus allowing fusion of datasets.

XYZ0987 Fused Dataset

Data source A

NRIC	Research ID
S1234567A	R1
S2345678A	R2

M1: Mapping table

Data source B

NRIC	Research ID
S3456789A	R3
S4567890A	R4

M2: Mapping table

First-level **hash is applied** to the mapping tables (NRIC/FIN and Research ID) **by the data contributors** before going into TRUST.
E.g., NRIC S1234567A -> ABC123
Research ID R1 -> TOK321

Other Data Source, e.g., HPB or Clinical Data

NRIC	Data1
S1234567A	…
S2345678A	…

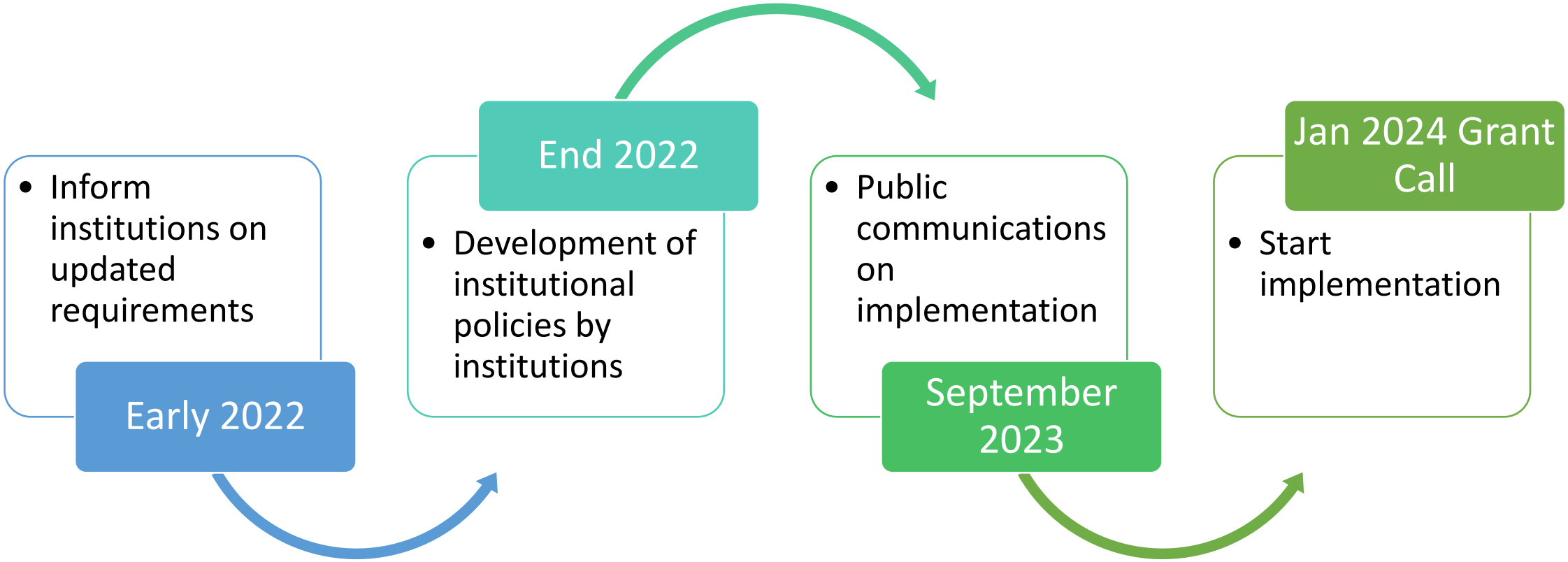
Raw data

First-level hash

Research ID	Data1
ABC123	…
ABC456	…

C1: De-identified data

Implementation Timeline



Annex

Data Sharing Plan Template

DATA SHARING PLAN TEMPLATE

The Data Sharing Plan shall be submitted in accordance with the template set out below. The notes in italics under each section of the template provide further context and guidance on how the section should be completed. Please refer to Annex B (Explanatory Notes to the Data Sharing Plan Template) of the Instructions to Applicants for further information.

1. Description of Research and Research Data

1.1 Brief Description of the Research Study

Briefly describe the nature of the research study (in not more than 3 lines)

1.2 Brief Description of the Research Data and Analysis to be Undertaken

Briefly describe the research data that will be collected/generated, and the plans for/scope of analysis to be undertaken.

2. Restriction on Data Sharing

You should share all the Final Research Data under the Framework within the stated timeframe under the Framework. Should you anticipate that this cannot be done, please state: -

(a) the scope of the Final Research Data that you will likely not be sharing under the Framework, or likely not be sharing under the Framework within the stated timeframe in the Framework;
(b) explain why that Final Research Data likely cannot be shared under the Framework or likely cannot be shared under the Framework within the stated timeframe in the Framework. Only strong justifications will be accepted; and
(c) include a proposal on alternative means in which that Final Research Data can be shared e.g. by anonymising or aggregating that Final Research Data, obtaining the relevant consent to disclose that Final Research Data; obtaining the relevant copyright permissions or sharing that Final Research Data at a later, more appropriate time (in which case, an extension of time should be sought pursuant to the Framework).

There is no need to fill in this section, if you are sharing all the Final Research Data under the Framework within the stated timeframe in the Framework.

*"**Final Research Data**" refers to, in relation to the Research, recorded factual material commonly accepted in the scientific community as necessary to document and support research findings, regardless of whether the data is used to support publications. Final Research Data that are clinical in nature shall include the human subject's basic demographics (including but not limited to the human subject's age, gender and race), disease and condition (i.e. the diagnosis), unless the information is not collected in the Research. Final Research Data shall not include laboratory notebooks, preliminary analyses, drafts of scientific papers, plans for future research, peer review reports, communications with colleagues, or physical objects such as gels or laboratory specimens.*

3. Data Use Limitation

Indicate any anticipated limitations to using the Final Research Data shared under the Framework. Such limitations should be minimised where possible and be substantiated with valid reasons e.g. where the limitation is due to limited consent obtained for the use and disclosure of the Final Research Data. Examples of such limitations include the use of the Final Research Data by non-profit organisation only, the use of the Final Research Data for health and/or biomedical research only, or the use of the Final Research Data for research related to a specific disease only.

4. Institutional Review Board (IRB) Assurance (For a Research Study that involves human subjects)

For research study that involves human subjects, please state whether an IRB has reviewed and endorsed the Data Sharing Plan.

Select Yes/No

5. Lead Principal Investigator's Undertaking

The Lead Principal Investigator shall undertake that the Data Sharing Plan is (i) in accordance with the Framework, and your institutional data sharing/management policy; and (ii) in compliance with all applicable legislation.

Please tick the box to undertake the following.

☐ The submitted data sharing plan is in accordance with the Research Data Governance and Sharing Framework, the Institutional Data Sharing/Management Policy, and applicable legislation(s).

Date: