Foundation Project Final Report

Introduction: This paper presents the IESO recommendations resulting from the Foundation Project. It is organized into the following sections.

- Purpose and Rationale of the Project
- Process Utilized
- Pre-implementation Issues Remaining
- Recommendations
- Next Steps
- Appendix of Final Recommendations

Purpose and Rationale of the Project: Ontario has made a significant investment in smart meters and in the central repository, the Meter Data Management Repository (MDM/R), resulting in high quality, consistent residential and small commercial electricity consumption data. The MDM/R is currently processing and managing the smart meter data to support local distribution companies' (LDCs) billing of electricity customers on time-of-use rates[1].

There is potential for significant improvement in harnessing the value of the MDM/R data set for designing conservation and demand response programs, system planning, policy development, academic research and to support innovation in Ontario.

The structure of the Ontario electric utility industry differs from many other North American jurisdictions of comparable or larger size. Currently, Ontario has to pool information from over 70 distribution utilities to assemble a provincial data set, while in jurisdictions with one large utility a jurisdiction-wide data set is automatically created. In Ontario, capturing the value of our amalgamated data set will involve successfully implementing the two items set out below:

- ensuring that the consumption data sent to the MDM/R includes geo-location, customer identifier or other relevant information to capture the analysis value of the data set; and,
- developing a framework or protocol that governs access to data from the MDM/R and/or the MDM/R Data Mart[2], and builds in "Privacy by Design".

The scope of the Foundation Project is to develop recommendations for each of these requirements. Any implementation of the recommendations would be undertaken as part of a subsequent initiative. For further details on the Foundation Project please refer to the Independent Electricity System Operator (IESO) stakeholder engagement website: http://www.ieso.ca/Pages/Participate/Stakeholder-Engagement/Foundation.aspx

[1] Excluding Toronto Hydro that is currently planning on integrating with the MDM/R in late 2017.

[2] The MDM/R Data Mart (scheduled for operation in the fourth quarter 2015) will contain a copy of the smart metering data in the MDM/R and is designed to service the increasing demand for the MDM/R data.

Process Utilized: This project was launched in early 2015 and a consultative approach was adopted to develop the framework. The Foundation Project's stakeholder engagement plan invited stakeholders to participate in the Foundation Project by providing input into defining the information to be added to the MDM/R to enable analysis and developing rules and protocols to provide access to such information. As part of this engagement, a working group of subject matter experts[3], the Foundation Working Group (FWG), including the Information and Privacy Commissioner (IPC) as an active observer, was convened to provide advice and guidance to the IESO.

The methods used for stakeholder engagement included in-person meetings (two of the five working group sessions were open to the public) and written feedback. The Terms of Reference adopted by the FWG provided that there would be no voting and no requirement to achieve consensus; however, where consensus was not achieved, the different views would be recorded. To the extent possible, consensus was sought on recommendations within the working group. In some topics areas, e.g. access to personal information, consensus could not be achieved and the diversity of views are reflected in this document. Therefore, the recommendations presented in this document are IESO recommendations developed with input from the FWG.

The SME Steering Committee was also periodically updated on the status of the Foundation Project. The recommendations were also presented to the IESO Stakeholder Advisory Committee for input and posted to the IESO public website for comment. In addition, the IESO undertook independent research.

Pre-implementation Issues Remaining: This document provides high-level recommendations to enhance the value of the data currently held in the MDM/R. No decision has yet been made on implementation. Prior to implementation, a number of matters would need to be addressed, including:

1. FIPPA compliance: the *Freedom of Information and Protection of Privacy Act* (FIPPA) sets out rules applicable to public sector entities regarding the collection, use and disclosure of personal information. Electricity consumption information when paired with address information may be personal information. As such, prior to collecting, using and disclosing this information, the IESO would need to ensure compliance with FIPPA requirements. The IESO would work closely with the Ministry of Energy and the IPC to address FIPPA requirements.
2. Cost estimates: this document sets out a number of options for third party access (although implementation is outside of the scope of the Foundation Project). Prior to

[3] The members of the Foundation Working Group can be viewed at the following link: http://www.ieso.ca/Pages/Participate/Stakeholder-Engagement/Foundation.aspx

implementation, it would be necessary to develop cost estimates and consider user/implementer needs, to assess which options, if any, to carry forward.

3. Data Matching: in the course of the FWG discussions, it was made clear that significant value comes from the ability to match electricity data with other data sets at the source level, prior to de-identification. If it is not possible to share personal information broadly, then FWG members felt that there would be value in having an entity that could match electricity data with other data sets and release the de-identified results. Prior to implementation, it should be determined what organization would be best suited to undertake the data matching service.

Recommendations:

- To ensure that the electricity consumption data can be meaningfully analyzed, at a minimum premise address information and premise occupant change occurrences must be added to the data set. In the absence of a premise street address, providing global positioning system (GPS) coordinates, if available, would be a suitable substitute.
    - Rationale: This approach builds on data already collected by LDCs and is expected to minimize administrative burden, while allowing data users to usefully segment the data and match it with other data sets to expand insights.
- Provide a framework for third party access to personal information and to personal information that has been de-identified to render it non-personal. Within the third party access framework, limit access to personal information to federal, provincial and municipal governments for purposes related to energy planning, policy development and other uses, as authorized by applicable legislation, regulation, licences or directives. Note that consensus was not achieved on this recommendation. See the full recommendation on third party access in the Appendix for a description of the different positions.
    - Rationale: The release of personal information is a complex matter. LDCs have expressed concerns about the release of personal information. Balancing the LDCs' concerns with the benefits to third parties who may have access to personal information, we recommend limiting access to personal information to government bodies, to put to use for the public good, under strict terms of use to minimize privacy risk. Such access requests and resulting action would be communicated to the LDCs whose customers' personal data is being accessed.
- Adopt a robust de-identification approach to ensure that privacy is protected.
    - Rationale: Under FIPPA, a custodian of personal information is required to maintain that data in a secure fashion to protect individual privacy. De-

identification of data is a recognized approach for protecting privacy and making data available for analysis.

Next Steps:

- Discuss the recommendations, and the possibility of their implementation, with the Ministry of Energy. The government is investigating the implementation of a MDM/R Data Access Platform or MDAP, a separate project, but having some common elements including the need to address privacy matters.
- Continue to work with customers, LDCs, the Electricity Distributors Association (EDA), government, the IPC and other stakeholders to expand access to energy consumption and other data, enabling, among other things, innovative energy management and the evolution of better data standards.

Appendix of Final Foundation Project Recommendations

This appendix contains the complete recommendations resulting from the Foundation Project. They are:

- Address and Occupant Change Information in the MDM/R
- Framework for Third Party Access
- De-identifying Information for Disclosure to Third Parties

The Foundation Working Group (FWG) engaged in discussions to determine the minimum set of information needed to enable the enhanced use of energy consumption data in the Meter Data Management and Repository (MDM/R), without imposing too burdensome and costly an implementation on the local distribution companies (LDCs) and the Smart Metering Entity (SME). Please refer to the session meeting summaries for these discussions. The following are the Independent Electricity System Operator's (IESO) recommendations, with input from the members of the FWG, for including Service Delivery Point (SDP) address and occupant change information in the MDM/R. Any subsequent implementation of these recommendations would be subject to compliance with existing privacy legislation and would include the participation of both the SME Steering Committee and the MDM/R Technical Panel.

GEO-LOCATION INFORMATION[4]

SDP (or Premise) Address Structure

Arising out of the Foundation Project is the recommendation that LDCs be required to provide accurate premise address information to the MDM/R using the existing "Premise Address", "City", "Province" and "Postal Code" fields. The Premise Address field would contain the components of street name (including, where applicable, street direction), street number and unit number, where they exist, ordered in a pre-specified sequence. The Premise Address field is a single alphanumeric field of up to 100 characters, for which the following structural options would be available to LDCs:

- Send the address as a single field using a predefined delimiter to separate the distinct components
- Send the address as a single field that is "free form" i.e. does not necessarily use a delimiter to separate address components.

[4] In addition to address and occupant change information, many other geo-locational elements were identified by the FWG as being useful to enhance the value of the MDM/R data; however, based on the established criteria for consideration within the Foundation Project, they were found to be beyond its scope. All these elements have been communicated to the MDM/R Data Access Platform (MDAP) project for their consideration.

SDP Address Standard

Also resulting from the Foundation Project is the recommendation that Canada Post be used as the address standard as it is the most widely accepted standard in Ontario.

It is recommended that, if necessary, the SME use post processing to parse address data provided to the MDM/R into a standard format to facilitate mapping using address information.

Global Positioning System (GPS) Coordinates

To the extent LDCs have collected GPS coordinates, (e.g. latitude and longitude) of their SDP addresses it is recommended that, in the absence of a premise street address, providing this information to the MDM/R would be a suitable substitute.

OCCUPANT CHANGE

It is recommended that LDCs provide occupant change information to the MDM/R using one of the following approaches:

- Send changes in the SDP to the "Account ID" relationship within the MDM/R that indicates an occupant change in the premise (existing MDM/R functionality already utilized by some LDCs).
- Create a new "date effective premise occupant change" field in the MDM/R for LDCs to populate (new MDM/R functionality).

It was also noted that for some LDCs, other information already in the MDM/R might be usable to determine an occupant change at a premise. It is further recommended that the SME and LDCs should explore this possibility further to potentially reduce the implementation cost of providing occupant change information to the MDM/R.

This document sets out a high-level framework to provide third party access to de-identified data and to personal information (the "Framework"). The Framework would only apply to the entity that holds the personal information across all jurisdictions in Ontario; however, local distribution companies (LDCs) may have access to the same information in their own jurisdictions and may be able to share this information with their authorized agents. This Framework is designed to support future projects that would implement the access rights discussed herein. This Framework assumes that all data will be handled in accordance with relevant legislation (including privacy legislation) and government directives; however, the potential impact of the draft government Open Data Directive has not been considered in this recommendation. This document also assumes that a robust de-identification process would be part of any implementation.

For the purpose of this document, the Data Custodian is the Independent Electricity System Operator (IESO)/Smart Metering Entity (SME) and the source system to be accessed containing the requested electricity consumption data is either the Meter Data Management and Repository (MDM/R) or the MDM/R Data Mart. However, the general provisions of the Framework could be applicable to different data custodians or different source systems containing other data.

As referenced in this Framework:

- "Data Custodian" is the entity that holds the personal information in the MDM/R or MDM/R Data Mart across all jurisdictions in Ontario.
- "Requestors" include any organization or individual submitting a request for data.
- "Recipients" refer to any organization or individual who will gain access as a result of a request being fulfilled.
- "Individual Users" refer to authorized individuals who will gain access to the data within the Requestor, Recipient, Data Custodian or their authorized agents.

With respect to third party access to personal information, the Foundation Working Group (FWG) was unable to reach a consensus. The specifics of the different perspectives on this issue are identified below.

Background

In 2004 the government of Ontario launched the Smart Metering Initiative. This initiative involved the installation of smart meters in homes and small businesses across Ontario and the establishment of the MDM/R, managed by the SME. The SME was created by the *Electricity Act, 1998* and the IESO is designated as the SME by O. Reg. 393/07. The *Electricity Act, 1998* s. 53.7(3) sets out the objectives of the SME, one of which is:

> *To provide and promote non-discriminatory access, on appropriate terms and subject to any conditions in its licence relating to the protection of privacy, by distributors, retailers, the IESO and other persons*

Categories of Access

This Framework addresses four potential access scenarios:

1. Standard reports developed using de-identified electricity consumption data downloadable from a public website by any third party (de-identified);
2. Access requests for de-identified electricity consumption data, alone or matched with other data sets, that are fulfilled by the Data Custodian, and made available to authorized third parties (de-identified);
3. Access requests fulfilled by the Requester through the Data Custodian's system access portal or interface, set up to permit access to de-identified electricity consumption data by authorized third parties (de-identified); and,
4. Access requests for electricity consumption data including personal information, alone or matched with other data sets, that are fulfilled by the Data Custodian and made available to authorized third parties (personal information).

1. Downloadable Standard Reports from a Public Website (de-identified)

Standard reports derived from the data in the MDM/R could be created and made available on the IESO's website for download by any third party. This would be similar in concept to the electricity system data currently available on the IESO website: http://reports.ieso.ca/public/RealtimeMktPrice/PUB_RealtimeMktPrice.xml

Funding the costs of developing and maintaining downloadable standard reports would be defined as part of the implementation of such a facility. Options could include these costs being part of regular SME operations or a separate cost recovery mechanism.

These reports would contain only de-identified data and would be in a format that could be used and analyzed. In developing standard reports, the IESO would consider reports based on a

sample group to lower cost and with regional identifiers and sufficient temporal granularity to make them useful.

The terms and conditions of use governing third party access to this type of data may be similar to those already contained on the IESO website, and would comply with the IESO and SME licences and legislation. (See for example: http://www.ieso.ca/Pages/Terms-of-Use.aspx).

Use of the data contained in the standard reports would need to be limited to the purpose for which it was collected, and re-identification of the data would be prohibited.

2. Custom Data Requests Fulfilled by the Data Custodian (de-identified)

This category would be for custom requests for de-identified electricity consumption data, either on its own or paired with other data sets, and fulfilled by the Data Custodian. The results would be delivered to the requesting party via an agreed upon delivery channel. This service could include the Requester providing the Data Custodian with additional data sets, or requesting the Data Custodian provide other data sets it has available, to be matched with the electricity data and then de-identified. The Requestor would not have direct access to the source system of the data (e.g. the MDM/R) nor would they have access to personally identifiable information held in the MDM/R or MDM/R Data Mart.

The Requestor would execute an agreement with the Data Custodian that includes terms and conditions of use governing third party access to this type of data which could include provisions governing some of the same situations contemplated by the standard reports discussed above, as well as basic rules of access, including:

1) A registration process completed by the Requestor, and upon successful completion it would be authorized to submit request(s).
2) Submitted requests would provide the following information:
   a) Specific data being requested,
   b) Intended use of the information (the Requestor agreement would provide it could not be used for other purposes),
   c) Intended disclosure of the information (what other parties would have access to the data and at what level of granularity), and
   d) Other data sets to be matched with the requested data (if any), including the expiration date of the licence for these data source(s) (if applicable)[5], information as to the level of granularity at which the data is to be matched and acknowledgment of compliance with any data access rules for the other data sets.
3) A request fulfillment assessment process would consider the following constraints and prioritization criteria:

[5] Any work or de-identified information that is the output of this process will not necessarily expire with the source data license expiration.

a) Current volume of requests to be serviced,
b) Complexity and amount of requested information,
c) Availability of resources (systems and personnel) and associated costs,
d) Use of the information,
e) Any laws, regulations or directives involved, particularly if deadlines have been imposed,
f) Organization requesting information, and
g) The expiration of any licence for additional data provided.

4) The requesting party would agree to certain terms and conditions, which may include the following:
   a) Cost structure for services,
   b) Payment terms for services,
   c) Requestor has in place effective controls to protect the security and privacy of data that meet or exceed the controls used by the Data Custodian (which they would require of any other Recipient of the data),
   d) Ability for the Data Custodian to require the Requestor (and any other Recipient of the data) to be subjected to an independent audit and provide evidence to the Data Custodian to demonstrate the design and operating effectiveness of its security and privacy controls, and
5) Data would be de-identified following an appropriate de-identification process (see "Recommendations on De-identifying Information for Disclosure to Third Parties" in this Appendix) prior to being delivered to the requestor.

3. Data Access Requests fulfilled using a Portal or Interface (de-identified)

The broadest data access to third parties would be to allow them to retrieve or access de-identified data from the Data Custodian using a portal or other interface into the source system. In addition to the terms and conditions of use contained in the above types of access (modified as needed), the IESO would also need to consider the following processes, rules and requirements:

1) The Requestor would submit an application to be qualified to directly connect to the portal or other interface, specifically accepting any terms and conditions relating to use of the portal/interface (in addition to the data itself). The application would be assessed by the IESO and only approved requestors would be allowed access.
2) The Requestor would comply with the registration, connectivity and authentication processes and tests for direct access to the portal or other interface by each qualified organization and would take responsibility for its Individual Users.
3) Access would be for a limited, agreed upon duration, and the Requestor would agree to timely notification of the IESO of the removal of authorized Individual User's access privileges to the portal or interfaces in accordance with the established processes and procedures of the Data Custodian.

4. Personal Information Requests that are Fulfilled by the Data Custodian

This category focuses on access to electricity information that is personal information, on its own or paired with one or more other data sets.

Providing access to personal information to a third party, while protecting that personal information, involves many complicated issues and affects multiple organizations. Given the complexity of such access and the fact that many use cases can be satisfied with de-identified data, the IESO is not recommending broad access to personal information at this time. The IESO's recommendation is that access to personal information be limited to federal, provincial and municipal governments for uses related to energy planning, policy development and other related uses, as authorized by applicable legislation, regulation, licences, or directives. Such access requests and resulting action would also be communicated to the LDCs whose customers' personal information is being accessed. The recommendation reflects governments' need for data, their experience with handling sensitive information, as well as the fact that they are trusted counterparties and serve the public good.

The access framework in this scenario may be similar to that noted for de-identified data; however, additional provisions may be incorporated, including in the following areas: (1) determination about who can access the data (including personal information) and for what purpose, and (2) additional control requirements over the security and privacy of the personal information.

Note: Summary of FWG Positions on Recommendation
Members of the FWG have differing opinions on the IESO's recommendation that access to personal information be provided to federal, provincial and municipal governments.

Some FWG members support the recommendation and recognize the various government use cases for access to personal information. One use case that was discussed was the municipalities' need to match building/premise address information and energy consumption data to facilitate community energy planning, including the development of Municipal Energy Plans, to inform infrastructure development, energy conservation and demand response program design and policy development.

We heard from the local distribution company (LDC) members of the FWG that they have a responsibility to protect the privacy of their customers' personal information. Therefore, some LDCs would only disclose personal information in compliance with legislative and regulatory requirements, or with customer consent.

Finally, there were other members of the FWG who supported the recommendations, but wanted to expand access to personal information to all, including private sector entities.

Conclusion

While this paper has outlined a Framework and considerations for access, the IESO notes that any future implementation would need to consider requisite legislative, policy and process requirements, including for the data matching service, that define in greater detail the governance framework, accountability and controls, including specifics relating to third party access to personal information.

The Foundation Working Group (FWG) engaged in discussions around protecting privacy while enabling sharing of the Meter Data Management and Repository (MDM/R) data. With the recommendation to add geo-location information to the MDM/R, the resultant data in the MDM/R may be considered personal information. Therefore, a framework for de-identifying personal information was determined to be required in order to facilitate access to de-identified data while protecting privacy ("De-Identification Framework").

An expert panel presented and provided resources to the FWG in respect of the de-identification of personal information for disclosure to third parties to inform these discussions. The preponderance of the expertise and information provided by the subject matter experts comes from the medical field. The processes and techniques developed for de-identifying personal health information are considered best practices. The Independent Electricity System Operator's Foundation Project recommendations in this area are largely based on those practices of the medical field.

Data de-identification is not a static process, where a single set of procedures, techniques and tools is established and exactly repeated for every de-identification operation. Therefore, the following is a recommended framework, based on accepted principles and practices, for de-identifying information. Each request for de-identified data will be assessed on a case-by-case basis, taking into account the purpose of the disclosure, the Requestor (defined below), the intended Recipients (defined below), the data sets that will be matched[6], previous requests and the risks of re-identification, among other considerations.

The recommended De-Identification Framework will be formulated in the following sections. This De-identification Framework would need to be adhered to when undertaking data de-identification.

1. Definition: De-identified Information
2. De-identification Context and Assumptions
3. Principles for De-identification

[6] While it is recognized that matching MDM/R data with other data sets at the source level is an important component of enhancing the usability of the data, who and how any matching services would be provided would be determined as part of the implementation of the Foundation Project, should it go forward.

4. Requirements for a De-Identification Process
5. Steps for the Data De-Identification Process

1. Definition: De-identified Information

The definition of 'de-identified information' the FWG has adopted is based on the definition in *De-identified Protocols: Essential for Protecting Privacy* from the Privacy by Design framework.

> De-identified information is information that cannot be used to identify an individual, either directly or indirectly. Information is de-identified if it does not identify an individual, and it is not reasonably foreseeable in the circumstances that the information could be used, either alone or with other information, to identify an individual.[7]

2. De-Identification Context and Assumptions

The following context and assumptions are made for the De-Identification Framework:

- This De-identification Framework is focused on de-identifying personal information contained in the MDM/R and the MDM/R Data Mart, after first matching this personal information with data sets from other sources that would be provided by the requesting party or be available from the Data Custodian.
- All de-identification of the personal information will be done by the Data Custodian or its authorized agent. For the Foundation Project, the IESO, as the Smart Metering Entity, is the Data Custodian of the MDM/R and related MDM/R Data Mart data. As referenced in this De-identification Framework:
    - "Data Custodian" is the entity that holds the personal information in the MDM/R or MDM/R Data Mart across all jurisdictions in Ontario.
    - "Requestors" include any organization or individual submitting a request for de-identified data.
    - "Recipients" refer to any organization or individual who will gain access as a result of a request being fulfilled.

[7] Source: Cavoukian, Ann, Ph.D., and Khaled El Emam, Ph.D. "De-identification Protocols: Essential for Protecting Privacy." 25 June 2014. Web: https://www.privacybydesign.ca/content/uploads/2014/06/pbd-de-identifcation_essential.pdf

- "Individual Users" refer to authorized individuals who will gain access to the data within the Requestor, Recipient, Data Custodian or their authorized agents.

- It is assumed that the Data Custodian will have the authority to collect, use and disclose the inputs to the de-identified data (the underlying personal information). The Data Custodian will comply with applicable privacy legislation in collecting, using and disclosing personal information. As such, customer consent will not be required for fulfilling any of the requests for de-identified data.

- The Data Custodian will define the services and terms and conditions of service, including for fulfilling requests to match data sets, de-identify personal information and stipulating certain consequences for non-compliance with the terms and conditions.

- The data Requestor will comply with the appropriate terms and conditions, including protection of privacy, and in the event the data Requestor is permitted to disclose the provided de-identified information to other Recipients, it will include appropriate terms and conditions of service in its agreements with such Recipients.

- For the purposes of these recommendations around risk assessment, an example of a typical vulnerability, threat and adversary would be:

  - Vulnerability – e.g. inadequate controls for authorized Individual Users to log on to a system containing the de-identified and/or personal information.

  - Threat – e.g. authorized Individual User accessing de-identified and/or personal information in a manner that may compromise controls.

  - Adversary – e.g. unauthorized organizations or individuals seeking to compromise the integrity of the de-identified and/or personal information.

3. Principles for De-Identification

The following principles are recommended for the de-identification of personal information:

- Enable appropriate collection, use, storage, and disclosure of de-identified data while maintaining public trust, privacy, integrity, security, availability, and confidentiality.

- Restrict access to personal information and to de-identified data to authorized Individual Users within the Data Custodian, Requestor, Recipients and/or their authorized agent, for authorized purposes.

- Establish a documented, clear, transparent process for de-identification with objective measures of risk.

- Establish a documented governance structure with clear accountabilities of all parties involved, including consideration of requirements of those parties to fulfill those accountabilities.

- Manage risk, in accordance with an authorized framework and risk thresholds of providing data access. Minimize re-identification risks while meeting the legitimate purpose of disclosure and acknowledging that risks cannot be entirely eliminated.

- Leverage available standards and best practices for the data de-identification process and techniques, as they evolve. Leverage appropriate systems and competent personnel for the design, implementation and operation of the de-identification process.

4. Requirements for a De-Identification Process

The following defines some of the requirements that must be included as part of the data de-identification process for every request that is processed:

- Define the roles, responsibilities, and scope of the Data Custodian, Requestors and Recipients, and, if applicable, their authorized Individual Users and authorized agents.

- Define the rules and procedures for prioritizing, managing and fulfilling/declining and logging all de-identified data access requests.

- Define criteria for determining what is an appropriate use of the de-identified data.

- Define when a privacy impact assessment and/or a threat/risk assessment are/is required and associated disclosure requirements.

- Define a process for evaluating re-identification risks.

- Define comprehensive and enforceable Requestor-Data Custodian agreements to obligate Requestors to maintain the confidentiality of the information that they receive. The agreements should specify at a minimum that the Requestor receiving the de-identified data:

    - Must not re-identify, or attempt to re-identify, or allow to be re-identified, any personal information

    - Must not link any other data elements to the de-identified data without verifying that the personal information will remain de-identified

- Must have controls in place that provide reasonable assurance that only the Requestor's authorized Individual Users will have access to the de-identified data and that it will remain de-identified

- Must agree that all of the data Requestor's authorized agents and its authorized Individual Users with access to the de-identified data shall abide by the confidentiality and privacy terms and conditions of the data Requestor-Data Custodian agreement to safeguard the de-identified data.

- Periodically conduct a review of the process and methodology against current and foreseeable threats and known vulnerabilities. This review may include reviews of the design and operating effectiveness of the process, as well as embedding monitoring controls in the process.

5. Steps for the Data De-identification Process

The de-identification process must include, but is not limited to, the following steps for each data request or request category:

- Determine whether the proposed use is an acceptable use of the de-identified data for each data access request.

- Identify and classify personal information data fields that are direct identifiers and quasi-identifiers.

- Perform an assessment of risks, both internal (such as accidental release of data, inadvertent access by employees) and external (such as the risk of successful external hacking of systems and databases, and risk of unauthorized users being able to re-identify de-identified data).

    - Identify plausible adversaries and plausible threats on the data (personal and de-identified) in question

    - Determine re-identification risk threshold or tolerance

    - Evaluate re-identification risk including, but not limited to, the considerations below:

        - Consider what reasonable resources with reasonable ability could accomplish in attempting to re-identify the de-identified data

        - Consider not only what data is being provided but also other data available or that may become available in the foreseeable future that could be used to re-identify the de-identified data

    - Establish acceptable risk threshold based on the specific Requestor and Recipients of the de-identified data

- Perform data de-identification to mitigate risk, including:

    - Define a set of data transformations (e.g. generalization, perturbation, data masking) for de-identification to reduce the risk to an acceptable level

    - De-identify data by applying appropriate data transformations, considering current standards and best practices for de-identification techniques and potential types of risks

    - Verify that the original personal information has been properly de-identified.

- Perform residual risk assessment of the de-identified data, including:

    - Assess residual risk and data usability after application of risk mitigations, considering plausible adversaries and threats

        - Has the acceptable risk threshold established for the specific Requestor and Recipients of the data been met?

        - Document risk decision

    - Determine whether to fulfill or decline the request

        - Compare evaluated risk with the acceptable risk threshold

        - If decision is made to decline the request, provide an explanation of the reasons for declining to the Requestor

        - If decision is made to fulfill the request, provide a description to the Requestor of the de-identification method used to enable the Recipients to make accurate use of the de-identified data, without exposing specifics that could lead to re-identification.