**Data Access Protocol (Template)**

Where possible please specify DHDP best practices and standards

([**FAIR**](https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/information-management/guidance-assessing-readiness-manage-data-according-findable-accessible-interoperable-reusable-principles.html) **principles,** [**OMOP**](https://www.ohdsi.org/data-standardization/) **CDM,** [**Flower AI**](https://flower.ai/docs/framework/) **framework**).

|  |
| --- |
| **Data characteristics** |
| **Disease Domain** | e.g., oncology |
| **Type(s) of data** | e.g., clinical |
| **Data provider(s) being organizations where data is sourced.** | e.g., Hospital A |
| **Data source(s)** | e.g., EHR at Hospital A |
| **Dataset size (n)**Please include total and a breakdown by data type or disease domain if applicable. | e.g., 10 000 |
| **How is a data point defined?** | e.g., 1 patient = 1 row |
| **Report the proportion of data currently structured into the** [**OMOP Common Data Model (CDM).**](https://www.ohdsi.org/data-standardization/)If None, write “None”. | e.g., None |
| **Report the proportion of data currently de-identified.** If None, write “None”. | e.g., 100% de-identified dataset |
| **Does the project team attest to deidentifying data and structuring into the OMOP CDM so it can then be shared through the DHDP Platform?** | e.g., Yes |
| **What other standards are currently used?** Write "None" if no standards are used. | e.g., mCODE |
| **Is the data currently being shared?** If yes, how? If not, how will the data be accessed? | e.g., Yes, data sharing agreements are in place |
| **Name and title of point of contact for privacy and compliance at the data provider’s organization.** | e.g., Bob Smith, Chief Data Officer bobsmith@hotmail.ca |

|  |
| --- |
| **Data access and approval** |
| **Is there an existing Data Access Committee (DAC) in place?** |  |
| **If there is an existing DAC, detail the structure and process. If there is no existing DAC, how will data access requests be triaged and approved in an ethical manner?** |  |
| **What is the expected turnaround time?** |  |
| **Who will be responsible for the data provider persona in DHDP to grant access through the Platform?** |  |
| **What measures will be taken to avoid or address lengthy turnaround time (e.g., escalation pathway)?** |  |

|  |
| --- |
| **Intended Use of Data** |
| **Use case(s) description** |   |
| **Is this a federated learning use case?** |  |
| **If applicable, what type of machine learning model(s) does this use case focus on?**  |   |
| **What other tools and coding languages does the use case include?** |  |

|  |
| --- |
| **Regulatory requirements** |
| Outline regulatory requirements that must be fulfilled for DHDP deployment to occur at sites where data reside. This includes requirements for data sharing with the intended party (or parties) through the Platform. Note: DHDP software must be deployed proximal to data due to privacy-first federated architecture. |
| **List the institutional, provincial or territorial, and federal requirements** that must be met to execute the proposed project. Examples include REB-approval or federal requirements for devices capturing patient data. |  |
| **Multi-institution and cross-jurisdiction:** Specify if data will be shared across multiple institutions as well as multiple provinces and/or territories. If applicable, what are the key considerations when navigating regulatory requirements across the different jurisdictions? |  |

|  |
| --- |
| **Data Onboarding Requirements** |
| Outline the data onboarding activities required to de-identify data, map data to the OMOP CDM and acquire data into the Platform. See OMOP data standardization [resource](https://www.ohdsi.org/data-standardization/) and additional resources at <https://www.ohdsi.org/.> |
| Guiding questions* **Resource Allocation:** How will resources be allocated to support data transformation and workflow activities to make data interoperable and implement the use case(s)?
* **Data Quality**: Describe procedures for ensuring completeness, accuracy, timeliness, and validity of data. How will errors or missing values be handled?
* **Data Readiness:** What is the current baseline and what are specific requirements that must be met to execute the proposed use case(s) (e.g., OMOP data mapping)?
* **Risk Mitigation**: Outline any risks to timely data onboarding and propose mitigation strategies.
 |