Regulation Establishing Procedures and Charges for City-Supported Data Integration Services

A. What is an Integrated Data System (IDS) and what is IDEA?

Philadelphia's Integrated Data System: Integrated Data for Evidence & Action (IDEA) integrates and stores individual-level data for City residents regarding their government contacts and services received from 2000 to the present. IDEA receives routine data feeds from data sources that cover the following government service areas: mental health, substance use, child welfare, juvenile justice, out-of-school time, early education, early intervention, lead exposure, Medicaid access, emergency shelter, housing, arrest, local incarceration, births, and deaths. IDEA is the central hub within City government that can support internal cross-systems data projects and partnerships with external institutions.

The goal of an Integrated Data System (IDS) is to bring together in one place the capacity to answer questions about the efficacy of the programs used to support citizens. Sharing data across systems holds out promise beyond what can be learned from any one data system. The integration of administrative data systems provides potentially even more compelling information on patterns of multi-system program use, costs and outcomes. Key uses include internal analytics, research, case coordination, development, and direct client outreach to benefit City clients.

B. Why does IDEA charge for support?

IDEA charges reimbursement costs for any proposal that requires our office to integrate, de-identify, and/or transform data from multiple sources **for use by an external entity** (e.g., researcher at a university). This policy includes any projects where a person or organization external to the City of Philadelphia will be receiving: (1) data that has been integrated and/or processed by IDEA or (2) data products created by IDEA (e.g., a table of aggregate counts).

IDEA frequently receives such external requests from academics and organizations. Researchers value the availability of the City's integrated data, the efficiency of receiving data from a single source, the formats in which they can request data, and the knowledge and capacity of the IDEA team to support both the administrative tasks to acquire data and knowledge of the data provided.

The City has determined that it is appropriate to seek reimbursement to help offset the costs and impact that external data requests typically have on IDEA's resources and staff time. Furthermore, IDEA is performing a function which is typically covered by the approved budget for external projects.

C. What does project support include?

Separate from any one specific project request, the IDEA infrastructure provides the following:

1. A unified data warehouse comprised of multiple City data sets
2. Matched records to link individuals to a City service encounter history
3. A robust process to support data requests and relationships with key, internal City agencies

For each request, IDEA also provides:

1. Project consultation to determine the project scope, data availability and formats, the administrative steps for approval, project timelines, and project deliverables.
2. Assistance with the development of documentation to request approval.
3. Routing of the proposal to the Philadelphia Department of Public Health Institutional Review Board, data contributing department for unit approval, and Law Department Privacy Review and execution of a Data License Agreement or contract.
4. Coordination with the Law Department to develop the Data License Agreement terms and project-specific requirements.
5. Development of the query specification to determine the data to be extracted, the necessary data transformation, and the format of the data to be transferred.
6. Assistance with establishing City sFTP credentials to transfer the data securely to the external requestor.
7. Performance of data extraction, transformation, and transfer of the data.
8. Consultation with the investigator team to understand and analyze the data provided.
9. Administrative support to renew project approvals as needed and to verify data destruction at the conclusion of the project.

D. Process for Making a Request for Cross-Systems Data Integration Through IDEA

1. Data requestors external to the City will first have to connect with a City sponsor (e.g., a City agency or department) that is willing to support the project. If the requestor is unsure of who to connect with, they can contact IDEA via email at idea@phila.gov.
2. The City sponsor will then submit **a proposal form** via IDEA's standard proposal intake portal. The form requires the City sponsor to provide basic information about the proposed project. Please note that the portal will prompt users to create an account if they do not already have one established. Users must enter an email address and create a password to establish an account.
3. Upon review of the proposal, IDEA will contact the City sponsor and data requestor to gather additional information about the project to better understand the scope of the request and whether necessary data is available to support the project.
4. The project proposal and additional information collected by IDEA will be shared with MDO leadership to determine whether IDEA will support the project. The following criteria will be applied:

   a. Verification of funding availability to pay for City time and efforts to meet the request
   b. Ability to negotiate access to external, non-City data source(s)
   c. Scale of impact on the community in terms of number of City residents likely impacted
   d. Number of total work hours required by IDEA staff to complete the request

5. Once a determination has been made, IDEA staff will reach out to the project requestor and City sponsor to let them know if IDEA will be supporting the project. If IDEA cannot support the project, the external requester and City sponsor are still able to pursue data requests directly through individual departments.
6. If IDEA is able to support the project, the external data requestor will be asked to complete a **Quote Form** so that IDEA can provide an accurate cost estimate for the data request.
7. IDEA will develop a quote for the data integration request based on resource allocation level. The project price quote will not be adjusted for increases in hourly rates during the life cycle of the

project.[1]

8. The City sponsor and the data requestor must sign the project quote form and return it to IDEA. This signed form indicates that the external data requesting entity agrees to pay the quoted amount for the services requested and that any additional requests will be subject to additional review and cost.

9. If IDEA believes that the project may qualify as research involving human subjects, the data requestor will need to submit an IRB proposal application to the [Philadelphia Department of Public Health's Institutional Review Board (PDPH IRB)](#) to determine if IRB review is required, in which case the requestor will need to submit their proposal and [additional required documents](#) to the PDPH IRB. IDEA must receive IRB conditional approval before any such research project may move forward.

10. After IDEA and City IRB processes are complete, projects will be submitted to the Law Department Privacy Review Committee for approval. If approved, the Philadelphia Law Department will negotiate and execute agreements between the requestor's organization and the City of Philadelphia.

E. How does IDEA determine project costs?

See Appendix A, entitled *Office of Integrated Data for Evidence & Action: Charges for Data Integration Services* for more information on the three levels for project charges and the criteria for determining the appropriate level.

In more general terms, the Executive Director for IDEA has determined the charging levels using three factors: 1) the steps to support a project, 2) the average number of hours and staff needed to support a project, and 3) a blended hourly wage rate for IDEA staff.

The pricing structure will be adjusted annually to reflect updates to the number of hours and staff needed to support a project or changes in the blended hourly rate for IDEA staff. The project price quote will not be adjusted for increases in hourly rates during the life cycle of the project.

F. Resource Allocation Levels

Based on the source(s) of the data being requested for a project, IDEA has identified three levels of resource allocation that correspond to the differential staff time required for IDEA to support these various types of work.

For example, projects that join new data extracts not currently available in the integrated data system are more labor- and cost-intensive than projects that utilize cross-systems data currently integrated into IDEA's data warehouse. Accordingly, the reimbursement cost charged for projects that access and integrate new data sources will be higher than projects that rely on existing data sources.

IDEA will determine reimbursement charges based on whether a project falls into one of the following three levels of resource allocation, with Level 1 projects being the least resource intensive and Level 3 projects requiring the most time and support (see Cost Schedule):

Level 1: All requested data are currently available within the IDEA data warehouse.
Level 2: All requested data are owned by the City but at least one data source is not currently housed within the IDEA data warehouse.
Level 3: At least one requested data source requires that IDEA access data from a non-City (external) entity.

G. Cost Schedule (Determined by IDEA)

Level 1: All data required for the requested data products are currently available to IDEA.

Cost: $6,880.00

Core Level 1 Activities:
- Consultation on project scope and data availability
- Support for project review
- Query specification development
- Data extraction, transformation, and transfer
- Support and consultation on data result set

Level 2: All data required for the requested data products are owned by the City, but IDEA does not currently have access to all the data necessary to execute the request.

Cost: $10,320.00

Core Level 2 Activities (in addition to all activities specified in Level 1):
- Management of incoming data extracts
- Interpretation of data organization and data business rules
- Data cleaning
- Data matching

Level 3: The requested data products require access to at least one data source from a non-City entity.

Cost: $20,210.00

Core Level 3 activities (in addition to all activities specified in Level 1 and Level 2):
- External data negotiation

Costs for Subsequent Data Runs
Note: If multiple runs of data are required for projects (e.g., the data product must be generated by IDEA three times within six months), the following additional charges will apply:

Cost of additional data runs for a Level 1 or Level 2 project: $2,000.00 per additional data run.
Cost of additional data runs for a Level 3 project: $4,000.00 per additional data run.

---

[1] The resources required to perform work at each level are determined based on estimated hours and a blended wage rate for staff involved for each activity described.