



Advancing Early Childhood Policy and Programs
through Strategic Use of Integrated Data

ecdataworks.org

Recommendations to the San Francisco Office of Early Care and Education for Establishing Data Governance

August 2018

Prepared by

Missy Coffey, Ph.D.

Phil Sirinides, Ph.D.

Contents

Introduction to the Data Integration Roadmap	1
Why does San Francisco Need an Early Childhood Integrated Data System.....	2
Creating and Using Learning Agenda.....	4
SF OECE Integrated Data.....	4
San Francisco ECE Data Governance: Overview.....	6
Establishing Data Governance in San Francisco Office of Early Care and Education.....	8
Data Governance Next Steps for the San Francisco	14
San Francisco Needs for Analytic Tools and Data Reporting	17
Summary of Phases.....	18
San Francisco Data Governance Charter	1

Introduction to the Data Integration Roadmap

San Francisco's Office of Early Care and Education has a significant opportunity to develop and deploy a new suite of tools and strategies for utilizing data. The long-term objective of this data effort is to maximize the impact of public funding of ECE efforts, especially positive outcomes for vulnerable children. This work is a continuation of multi-year efforts focused on coordinating and aligning systems and programs while improving system-level efficiencies. The goals of these efforts include increasing quality of programs, access and ease of utilization, and professional capacity in order to positively impact providers, children and families, and all system partners and stakeholders. In the shorter term, OECE and partners can improve its management of information across multiple agencies in differing formats, and integrate data into daily operations of the citywide ECE system.

This document is the natural next step to the planning effort performed for OECE's initiative to integrate and leverage administrative data. OECE's agency leaders and key stakeholders have participated in this effort and are prepared to drive these work plans forward. In sum, OECE is well positioned to take a material leap forward in improving early learning data, which in turn will positively benefit early learning and child outcomes across San Francisco. However, as we describe in this report to accomplish these goals, OECE and contracted partners must 1) integrate data across systems, 2) build analytic capacity for data analysis, and 3) establish data governance to ensure sustainability of the integrated data system. Specifically, the creation and implementation of 1) a unified system for collecting and managing production data, 2) a data warehouse with tools for reporting and analysis; along with 3) data governance structures and processes are all essential to the success and sustainability of OECE services, and are the focus of this report.

To implement this vision San Francisco's OECE has worked with report authors, Drs. Missy Coffey and Phil Sirinides, who are national experts in Early Childhood Integrated Data Systems and co-lead the ECDataWorks center at the University of Pennsylvania. We have a recommended set of activities and prioritized them into specific next steps and a timeline for OECE and partners. ECDataWorks encourages the development of a learning agenda to articulate the information needs and phased priorities. To accomplish these goals we outline a set of technology phases that align to the needs of San Francisco's early care and education system. The first and most immediate objective for OECE is to obtain access to all data collected through their contracted partners to begin developing the infrastructure and processes to effectively steward the public resources and improve child outcomes. To inform the decisions that will be made throughout the development and implementation of the integrated data system, we recommend a series of actions to build an established data governance structure and process. To ensure the data is useful and provided in a timely fashion to OECE and its partners, we provide a set of recommendations on building analytic tools. At the end of the roadmap report we provide a summary of the timeline that encompasses short, medium, and long-term objectives to divide

the amount of work into more manageable parts. We describe what OECE needs to do within each phase.

Why does San Francisco Need an Early Childhood Integrated Data System

San Francisco has always been dedicated to finding innovative ways to support the children and families of San Francisco. In the last decade the city has approved policies that continue to support a comprehensive system of care for children and families in San Francisco and that seeks to improve access to high-quality care through coordinated services. As outlined in the *San Francisco Citywide Plan for Early Care and Education*, the development of the Office of Early Care and Education was the first step in building a coordinated model.

In 2013, the late Mayor Edwin Lee established the OECE and subsequent legislation outlined duties and functions of this office (Ordinance No. 261-14). The intent was to maximize benefit for children and families by centralizing processes which were managed by separate agencies. In order to fulfil these responsibilities, OECE needs to work to create an integrated data system that will allow them to “align and coordinate the City’s efforts to provide early care and education” across multiple funding streams. This data system will provide the necessary information to “improve access to high-quality programs, strengthen the workforce, and build the ECE system capacity.” Without data provided by an integrated data system, decisions about programs and services for SF children and families will continue to be fragmented and require increasingly more manual time from data analysts to pull data together to inform decisions.

An independent Comprehensive Fiscal Analysis completed for OECE in 2016 made recommendations for increased efficiencies that require the agency to develop an integrated data system. Integration of data is needed to reduce the workload of staff and time needed to respond to questions, allowing more transparency in the use of public funds to support early care and education. This current report represents the second set of external advisors who have observed a lack of capacity by OECE for data collection and use. Further delays in developing linked data systems will prevent the agency from successfully performing its duties and functions. We urge this office to begin rapid implementation of an integrated data system and also to establish data governance.

To better understand the information needs and current challenges with data, beginning in 2017 OECE conducted a study to identify stakeholders perspectives and there were two key challenges identified, (1) Existing data systems are fragmented, especially in terms of enrollment and CQI data, “[We] only know about ages served [in the programs they work with], not income levels or other needs...If I know there are 3 children in a program that all have open CPS cases, I might be able to help that FCC program think about family engagement in a different way. Without that information, we are only responding to what programs ask us for in terms of assistance” (FCCQN); (2) A lot of duplication currently exists, which creates burdens throughout the system

but particularly on programs, “It would be ideal if Wu Yee, Children’s Council and OECE were all using the same info and same system, so we don’t have to spend time trying to figure out what’s going on. It would help programs too. They currently have so many log ins. They have to send [copies of their license or bank info] to three different departments within one organization... Why not have it centralized?” (Nutrition) and “How can we make it easier on providers – only essential data from them and what is the easiest way to get it” (Subsidy staff).

Yet, the stakeholders also had ideas for how the integrated data system should work and three themes were identified: (1) OECE and partner agencies responsibly share data to improve efficiencies and outcomes, “We need to share data across data systems, and we need data standards for data sharing” (First 5) and “It is great to collect data, but is that data promoting children’s outcomes? How are you proving your investment is making an impact on developmental outcomes for children? We don’t have that type of data yet” (FCCQN) (2) Integrate functionality where feasible, “Ideally we would want a system with all of the functionality of CC3 and all of the provider referral features from NDS” (Subsidy staff) and (3) Specific requests for functionality in any new data system(s) such as a “direct query database for analysis” (Subsidy staff), “it would be great if those consultants could track how many of their providers did a ASQ and the follow up for each kid. It would enhance their consultant coaching efforts if they could say, ‘Let’s see how many kids you had concerns about. Let’s see who we can refer those children to’” (First 5), and “some sort of connection to state licensing. Some sort of flag if there is a violation. Right now, you have to log on to sign up for each individual program” (FCCQN). These are all stakeholder comments that OECE and partner leadership are considering as the SF ECE integrated data system is developed.

To fulfill its responsibilities, OECE identified the need for an integrated data system and has developed the following, provisional vision statement for the SF ECE data system:

The San Francisco ECE integrated data system will be a well-coordinated, integrated, and efficient system that allows families, programs, OECE, and its partners to use high quality, real-time data to promote favorable outcomes for children, age birth to five. The goal is not only to build an integrated data system, but to use it to inform decisions that support children and families in San Francisco.

The long-term goal is for OECE and its contracted partners to participate in cross-city department data sharing. Sharing data across San Francisco public agencies will allow OECE and contracted partners to support ECE program development and integration with the overall wellbeing and trajectory over time of families who touch multiple city systems (e.g. child welfare, dept of public health clinics/ER/mental health, CalWORKS, WIC, Medi-Cal, food

stamps, publicly funded Family Resource Centers, libraries, recreation and parks programs, job placement programs, homeless services).

Creating and Using Learning Agenda

This section outlines the definition of a learning agenda and the reason San Francisco should work to create a learning agenda as part of its data system integration effort.

How many children in San Francisco receive high quality early learning services? This information is not available. Of all kids who are most at risk – are we reaching them with high quality programs? This information is not available. How are children in high quality early learning programs doing compared to their peers? This information is not available. For San Francisco early childhood policy and programing to be effective and sustainable, OECE requires new solutions to leverage early childhood data for program advocacy and outcomes.

At the core of these questions is an interest, by the OECE and its partners, in building a complete picture of families and their experiences when a child is enrolled in an early learning program. In many ways, San Francisco has done much to increase access to quality, and to engage families in a conversation about the importance of quality in early care and education. There are several internal and external partners helping to advance a systemic approach to family engagement in San Francisco that goes beyond individual policies or activities.

A collaborative effort to create a learning agenda for children and families is the natural next step to recent planning efforts by OECE and key stakeholders as they work toward improved data practices. A learning agenda can support evidence-based policies that positively benefit early learning and child outcomes across San Francisco. A learning agenda is a set of prioritized questions which reflect the current information needs of stakeholders. Questions for a learning agenda should address an ongoing program need and offer clear decision points when answered. System developers can then prioritize enhancements to data systems and reports to make data available as needed.

Creating and using a learning agenda is important because data collection is not enough to support effective data use. Effective and sustainable data use depends on broader agency commitment to evidence-based decision-making and will require a combination of technical and non-technical factors. In order to implement a learning agenda, it will be important to build capacity in three important areas related to early childhood data: technical capacity, analytic capacity, and data governance. To move forward in creating a learning agenda, SF's OECE is working with Drs. Phil Sirinides and Missy Coffey, to develop a roadmap with recommendations for specific next steps and a timeline to create an integrated data system, to build analytic capacity, and to establish data governance. Each are described below.

SF OECE Integrated Data

In this section we provide recommendations of the integration phases based on conversations at the June meeting and further conversations and analysis of data systems supporting San Francisco ECE.

The first and most immediate objective for OECE is to obtain access to all data collected through their contracted partners to begin developing the infrastructure and processes to effectively steward the public resources and improve child outcomes. While disparate data systems can be sufficient for local monitoring and reporting, an integrated data approach is necessary for enhancing strategic planning, guiding technical assistance, ensuring program integrity, and fostering system quality improvement. This is because decentralized systems with different data standards lead to disparate business practices and ultimately service barriers as business partners behave autonomously. Without a single integrated system, partners are unable to share real time information that supports effective program coordination and management. A single information system with centralization of program data will enable consistent application of policies, notifications and business practices. Importantly, a centralized data system is necessary for answering questions reflected in a robust shared learning agenda.

We recommend, by the end of the calendar year 2018, OECE refine the child and program data systems it uses and that are used by its partners. Within three years, OECE should fully integrate data systems, and be able to demonstrate that services are helping families and children at a reasonable cost. We recommend the following phased development to leverage the existing integration and build toward the full integration across data systems to inform the OECE learning agenda.

Phase 1: Family Matching (2018)

Phase 1 is close to being completed through deployment of new wait list with family matching to available services. The new wait list system, Early Learning SF, is planned to launch its first phase of development in October 2018. By gathering more robust information from families and from ECE programs, the system's algorithm will make better and faster matches, which should help vacancies be filled more quickly. Also, instead of matching one family at a time to once vacancy, the system can match families with multiple available programs that meet their parameters, and match programs with multiple potentially interested families. The new wait list will be available in English, Spanish and Chinese.

Phase 2: Program and Child Data Integration (2020)

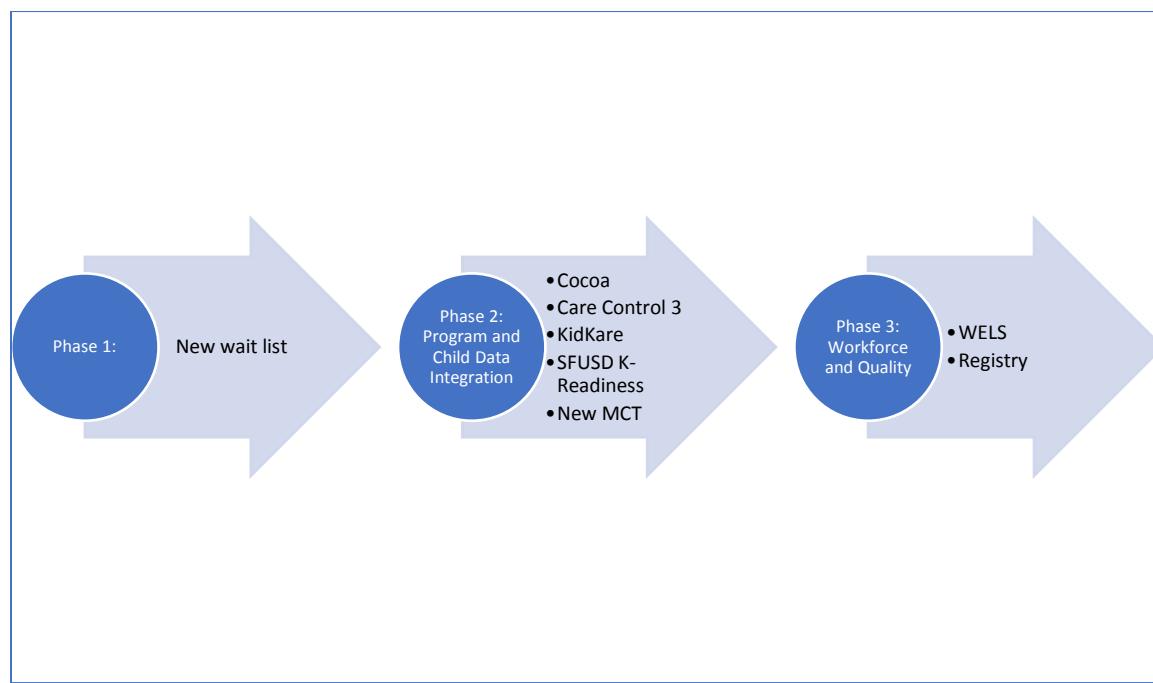
Phase 2 is focused on OECE management of information currently in Cocoa, two versions of Care Control 3 (one at Wu Yee and at Children's Council), and KidKare (both at Wu Yee and at Children's Council), such as enrollment, attendance, and payments. This phase should also include child outcome data such as the kindergarten readiness indicators from SFUSD. In addition to increased data integrity, OECE will be able to calculate distinct count of children enrolled in real-time. Phase 2 is planned to be completed by Spring 2020.

At the end of this phase OECE and the partners will be able to answer critical questions from the learning agenda and take action to inform decisions made by OECE and partners to provide quality services to the children and families of SF.

Phase 3: Program Quality and Workforce Data Integration (2021)

Full integration of OECE data comprise phase 3 to be completed in 2021 at the earliest. This will bring in quality and workforce data currently in WELS and the California Early Care and Education Workforce Registry. The long term strategy for cross-city agency integration (for example, linking students in San Francisco Unified School District's third grade academic and other school success indicators with past participation in OECE-funded services or examining the trends of families accessing both OECE-funded and SF Department of Health-funded services) has no timeline established and will depend on established inter-agency data governance and the success of prior phases of work.

At the end of this phase OECE and partners will be able to learn from critical information relevant for continuously improving the quality of programs and helping to provide necessary resources to support workforce.



San Francisco ECE Data Governance: Overview

In this section we provide a summary of data governance and make recommendations for establishing San Francisco data governance with OECE and related partners.

To build a sustainable integrated data system, ECDataWorks is proposing that OECE create a data governance structure and processes across the agency and its partners. The goal of data

governance is to ensure that agency staff and partners all contribute to decisions about the integrated data system. The United States Agency for International Development (USAID) offers one of the best examples of a public organization attempting to improve its agency's effectiveness by becoming a more effective and collaborative learning organization. USAID's Collaborating, Learning, Adapting (CLA) Framework identifies enabling conditions of culture, processes, and resources which are integrated across all operations. Based on San Francisco's similar context of having many partners with different data needs and practices, we see data governance as a strategy for creating more cohesion across existing data systems and helping OECE and its partners answer key questions about how to best invest local resources for maximum impact.

According to the U.S. Department of Education, National Forum Guide on Education Statistics (2011) "Data governance is both an organizational process and a structure. It establishes responsibility for data, organizing program area staff to collaboratively and continuously improve data quality through the systematic creation and enforcement of policies, roles, responsibilities, and procedures. There can be program data governance (single data system) or integrated data governance (multiple data system contributors)."

Data governance specifically oversees how information assets can be used to help achieve the early childhood goals. To accomplish these goals, OECE needs to create a set of policies and a group(s) with authority to make decisions about the integrated data.

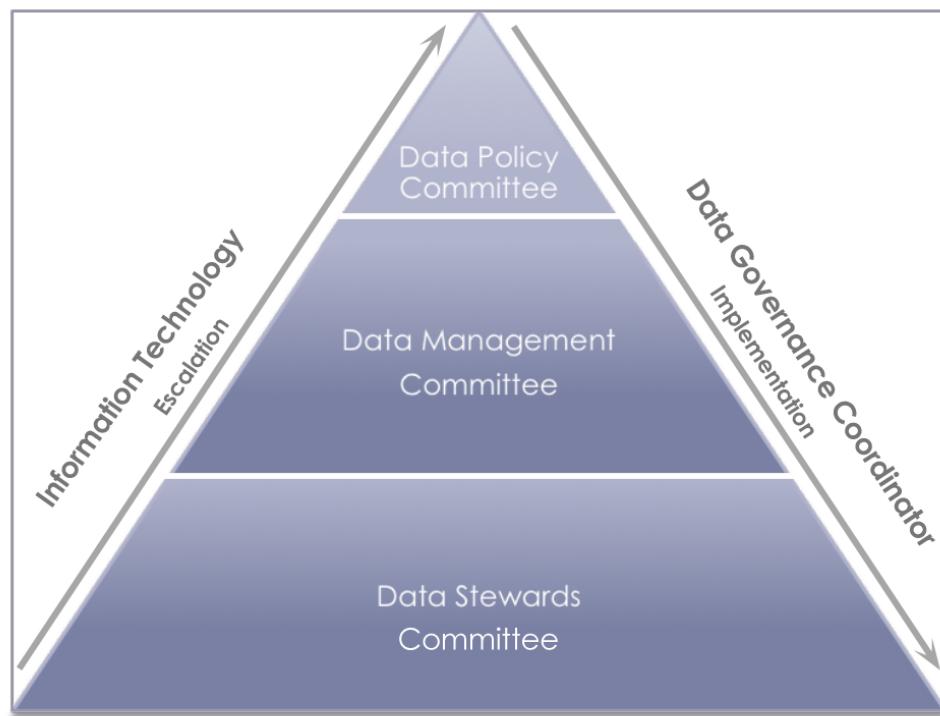
Typically data governance groups will work to establish a list of the policies that need to be created to manage and integrate the data in the following eight (8) areas.

1. Scope –What are the most important questions of the learning agenda? What data will be included/integrated?
2. Integration – How will data be brought together and normalized?
3. Refresh – How often and how will data be updated to inform the learning agenda?
4. Changes – What is the input process for changes needed to inform the learning agenda? What is the input process for changes to data standards and systems?
5. Access – Who will have what level of access to the data to inform the learning agenda?
6. Confidentiality- How will data be kept confidential and secure while using data to inform learning?
7. Requests – How will external requests for data be reviewed and responded to?
8. Releases – How will reports on insights based on the learning agenda be disseminated?

To address each of these critical data management topic areas, individuals that understand the data elements as well as those that have the authority to authorize changes or release of data need to be included in the overall structure of data governance. As outlined in Early Childhood Data Governance (Cochenour & Hebbeler, 2015), typical data governance structure consists of three groups: executive committee, data management committee, and data stewards (Figure 1).

Although it is not necessary to create new groups it is important to identify groups that will have the critical decisions and responsibilities.

Figure 1. Early Childhood Data Governance diagram from Early Childhood Governance: Choices and Consequences (Cochenour & Hebbeler, 2015)



Overall, data governance is important as SF OECE integrates data to inform decisions because once established it will:

- Prioritize the learning agenda and what data to integrate as you shift from individual collections to integrated early childhood system;
- Clearly delineate agency executive, and management, and data stewards roles to allow contributions and collaboration across organizations; and
- Provide opportunity to be more responsive to OECE and partner's information needs.

Establishing Data Governance in San Francisco Office of Early Care and Education

As a result of the meeting with SF OECE and partners agencies on June 22nd, 2018, national experts Drs. Coffey and Sirinides recommend establishing data governance to better understand where there is duplicate data being gathered and where data quality can be improved that allow the data to come together to make informed decisions. Data quality and duplication was the primary reason the participants in the meeting were interested in data governance. Second, there

was an interest in all representatives having a voice in the process to articulate vision, identify key policy questions, and make sure that their programmatic needs are met in the broader conversation across San Francisco.

Given OECE is hoping to integrate the following data systems in the initial three phases of this work, representatives working with the current data from each of the partner agencies and OECE should be included:

- New wait list system
- NDS R&R data system
- Cocoa
- Care Control 3
- KidKare
- WELS
- SFUSD K Readiness Indicator data
- CA ECE Registry

While OECE currently has a phased approach to refining the technology used across the SF ECE system, data governance should not be phased as it is a process that takes time. For example, although the California Department of Education has stated it plans to integrate the WELS (or another QRIS data system) and the CA ECE Registry in a few years, it is important to integrate the staff critical to those decisions and data from the beginning of the data governance process. This will ensure that they can contribute to the establishment of policies that will affect them at a later time.

To align to the goals of OECE we recommend the following structure and provide our rationale. Table 1 outlines the complete list of each committee and the representative role from OECE staff and partner agencies. The roles are mentioned intentionally to avoid the work depending on a specific individual. If an individual transitions to a new job then the person who is hired for the role will replace the former representative on the committee.

Table 1. Recommended San Francisco Office of Early Care and Education Data Governance

Data Governance Structure		Recommended Roles					Recommended frequency
		SF OECE	Children's Council	Wu Yee	SF USD	First 5 SF	
Executive Leadership committee	Executive Director	Executive Director	Executive Director	Early Education Director	Executive Director	Biannual or quarterly depending on the number of decisions being made	
	Deputy Director	Deputy Director	Chief Program Officer	Director of Systems and Operations	Senior Program Officer		
Data Management Committee	Finance/Subsidy	Fiscal Strategies Manager	Subsidy Manager R&R Manager	Subsidy Manager R&R Manager	EED Finance Manager	N/A	Every other month
	Quality Rating/WF	Quality/Workforce Analyst	FCCQN Manager	FCCQN Manager	Executive Director Program Quality, & Equity	QRIS Coordinator	
	Data/Research/IT	Data & Eval Manager	IT Director	Data Analyst	Supervisor of Analytics	Senior Program Advisor	
Data Stewards committee		Data & Eval Analyst	Data analyst	Data analyst- trying to hire/make a position	Data System Analyst	Database Analyst	None, they are to inform the representative of the data management committee

The Executive Leadership Committee. This is a group responsible for the final decisions about the data system and how data will be collected and used. This group should include executive leadership across all agencies partnering with OECE to provide services to children and families. The director of OECE referenced an existing group of leaders from Wu Yee and Children's Council and that others including San Francisco Unified School District and First 5 San Francisco should be included in meetings going forward. This entity was established to discuss key issues related to the system-level services delivered by Children's Council and Wu Yee. With the membership expanded as described above, data governance key issues will be added to their scope as they meet biannually, with a goal of moving toward quarterly meetings. The executive leadership committee is responsible for establishing and sustaining data governance. It does this by:

- appointing staff from each agency to participate in data governance;
- approving the data governance manual;
- creating the data management committee and appointing committee members;
- resolving issues escalated by the data management committee;
- approving data policies and major data-related decisions proposed by the data management committee; and
- holding all partner organizations accountable for participating in the data governance program and adhering to the data governance charter and manual.

Individual executive leadership committee members are responsible for:

- modeling and advocating data governance principles with across OECE and partner organizations;
- communicating with the data management committee members who report to them to prepare for each executive leadership committee discussion and decision;
- including data governance responsibilities in the performance expectations of their staff who are members of the data management committee and data stewards committee; and
- holding all their staff responsible for adhering to the data policies and processes established through data governance.

The Data Management Committee. This committee will be the most important for OECE and the partner organizations. The data management committee is responsible for making decisions regarding policy implementation and operations for specific functional areas and participating in decisions requiring system-wide policy and operations for the learning agenda. Members in individual organizations from Children's Council, First 5, OECE, San Francisco USD, and Wu Yee are responsible for ensuring that integrated data system applications that create, capture and update data elements, incorporate, edit and validation checks to protect data integrity. The data management committee makes policy recommendations to the executive leadership committee. Data management is chaired by the data governance coordinator. Overall, the program

management-level involvement in data governance is focused on establishing and maintaining policies and processes regarding data collection, quality, availability, and use. The data management committee does this by:

- establishing, documenting, and enforcing data standards, policies, and processes (e.g., data request policy and process);
- identifying, prioritizing, and supporting the resolution of critical data issues that affect the quality, availability, and use of data;
- establishing and monitoring the progress of workgroups of data stewards to research and propose recommendations for addressing prioritized data issues;
- escalating to the executive leadership committee controversial, high visibility, or high risk issues for resolution and policy recommendations for approval; and
- holding their program area staff accountable for participating in the data governance program and adhering to the data governance charter and manual

The data management committee is chaired by the data governance coordinator and is made up of organization managers who are responsible for program areas that require the collection, management, and reporting of data, as well as those with data-related responsibilities. Individual data management committee members are responsible for:

- modeling and advocating data governance principles with internal staff and other external stakeholders of their program area;
- attending data management committee meetings and participating in the decision-making process;
- ensuring that their program area data stewards and other staff adhere to the policies and decisions of the data management committee;
- sharing with the data management committee data issues, regulations, plans, and policies from their program area that may have an impact on other programs; and
communicating with their organization staff the activities and decisions of the data management committee.

Data Stewards. Given the staff and partners arrangements in SF OECE, the data stewards may not need to meet as a group, but they should be consulted by the data management/policy committee representatives on data elements and technical questions prior to changes being implemented. The data stewards are those that know the data and the quality of each piece of information. The knowledge they have is invaluable in the success of designing and using an integrated data system. Although they may not meet, they should have a clearly defined role and understand the responsibility they have to inform the data management/policy committee. Data Stewards are responsible for:

- determining how data are defined, collected, quality assured, and reported to meet the requirements and the agency's and external stakeholders' data use needs;

- informing calculation of metrics and determination of the source system of record for external reporting
 - reviewing and approving data releases
- communicating organizational needs for data privacy, security and archiving;
- documenting core processes regarding data collection, calculation, and reporting;
- ensuring metadata is documented and maintained;
- identifying critical data issues that impede data quality and use and participating in work groups to identify, propose, and implement resolutions;
- tracking federal, state, and other pending legislation or regulations involving data elements in his/her domain and communicating the potential impact;
- serving as a point of contact for data providers (e.g., local providers) for the given data domain;
- regularly evaluating and proposing process changes to improve data quality and/or agency efficiency;
- supporting their data management committee member by communicating data governance policies, processes and decisions to others within the program area;
- contributing information and expertise to support issues and discussions in data governance meetings.

Data Governance Coordinator. The data governance coordinator is a critical role to the implementation of data governance. In order for various decisions to both escalate and be implemented (as shown in figure 1) there needs to be an individual who coordinates and communicates across all committees. The Data Governance Coordinator has an agency-wide perspective and understands how data use should support and inform the agency's policies and programs. The data governance coordinator may have additional duties beyond data governance and may be an external partner, but should have adequate time to manage the data governance effort. In addition, the coordinator:

- facilitates the escalation and resolution of issues among the data governance groups;
- serves as a member of the data management committee;
- convenes working groups of data stewards to address critical data issues that span multiple organizations.

The knowledge, skills, and experience needed to be successful in the role of Data Governance Coordinator include:

- an understanding of how data use should support and inform education policies and programs, and an understanding of IT concepts and systems (but does not require an IT background);
- the ability to balance agency-wide goals and strategy with making and implementing specific decisions in alignment with the broader goals and strategy;

- project management knowledge and meeting facilitation experience;
- experience leading process development and revision;
- training or mentoring experience;
- strong interpersonal skills, including the ability to mediate multiple perspectives and come to a resolution;
- strong communication skills, especially in translating between program areas and technology;
- understanding of data quality principles and strategies; and
- solution oriented.

In addition to the data governance committees, other stakeholders will be engaged as stakeholders to inform the decision regarding the OECE integrate data system. Other organizations and roles that will be included in stakeholders:

- Family Child Care Representatives
- Child Care Providers Association
- Co-op association
- Title 5 Providers Group
- Head Starts/EHS

Stakeholders. The data governance body should be informed by stakeholders who are interested in the outcomes of the SF OECE integrated data system as well as being users of the data but are not data contributors themselves to the data system. Stakeholders could include R&R representatives who work with families and can represent the data needs of families, Family Child Care Association, Head Starts grantees, CPAC, Title V programs. The stakeholders can provide information that influences the direction, priorities, and implementation of the SF OECE. An intentional plan to gather regular input from various stakeholder groups should be created and shared with stakeholders to ensure they see how they can contribute to the development and implementation of the SF OECE integrated data system.

Data Governance Next Steps for the San Francisco

We recommend the following phased 18-month plan.

Within 1-3 months

- Executive leadership should meet to discuss the proposed approach and a data governance charter
- Compile the list of questions that OECE and the partners would like to address using an integrated data system (these should be questions that cannot be answered today without a great deal of time and effort) that build towards a learning agenda
- Develop one use case that demonstrates how to answer one of the questions using data and leading to action

- Identify who will play the role of the data governance coordinator
- Work with DG coordinator or external partner to facilitate:
 - Document the vision of integrated data (and within it data governance) for SF OECE, including: Why now? How does it help to accomplish the goal of integrated data? Why should others care? (*Much of this is captured in this roadmap and the data governance coordinator can work to create resources from the information provided here*)
 - Create a working definition of data governance for SF OECE
 - Articulate what data governance terms need to be defined
 - Identify the training needs of the various individuals involved
 - Train staff and partners on the implementation of data governance

How do you know if you have been successful?

- A data governance charter that documents the proposed structure and processes has been signed by executive leadership
- There is a list of questions that need to be prioritized
- A data governance coordinator is in place and planning for the rest of the phases (including the development of the data governance manual)
- OECE staff and partners can articulate the need/benefits for integrated data and data governance in SF OECE

In months 3-6

- Executive leadership begins to assign staff and identify key issues they would like to discuss regarding the data system
- Complete a needs assessment of how decisions related to data are currently made across OECE and partners. This information will help to inform what is currently working that can be built upon and what practices are going to be modified and better supported with data governance around the integrated data.
- Create the data governance manual that addresses the following eight (8) topic areas:
 1. Scope –What are the most important questions of the learning agenda? What data will be included/integrated?
 2. Integration – How will data be brought together and normalized?
 3. Refresh – How often and how will data be updated to inform the learning agenda?
 4. Changes – What is the input process for changes needed to inform the learning agenda? What is the input process for changes to data standards and systems?
 5. Access – Who will have what level of access to the data to inform the learning agenda?
 6. Confidentiality- How will data be kept confidential and secure while using data to inform learning?
 7. Requests – How will external requests for data be reviewed and responded to?

- 8. Releases – How will reports on insights based on the learning agenda be disseminated?
- Prioritize the list of questions and articulate the actions that would be taken if you had the answer to the question (if there is not actionable decision/change that can be made it should not be a priority question)
- The key structure (executive leadership committee, data management committee, and data stewards committee) is tested as it is executed, providing an ongoing feedback mechanism for both formal and informal feedback, and modified to include all the necessary staff and partners

How do you know if you have been successful?

- Creation of a learning agenda. Staff and partners can articulate the 3-5 questions that have been prioritize established across OECE and partners

In months 7-12

- Conduct a data discovery, including a review of the data dictionaries across all data systems that will be integrated

How do you know if you have been successful?

- Individual agencies are not making decisions about data integration to accomplish the learning agenda, but bringing the recommendations to the appropriate body to review and decide.

In months 12-18

- Review the learning agenda and reprioritize as needs change
- Implement data governance manual process (including the decision-making process)

How do you know if you have been successful?

- Individuals use terms such as “our data” not “my data” referencing the individual agencies that they represent

Overall, SF OECE and partners will know when data governance is established data governance when:

- The learning agenda established aligns to the needs of OECE and the partner organizations.
- Well-defined policies concerning seamless early childhood education (collaborate on developed key questions)
- Coordination between agencies and programs administering early childhood services and collecting data
- Improve the quality of data collected, reported, and used by multiple agencies and early childhood programs
- Reduce staff time and effort to

- Improve communication, collaboration, and relationships between OECE and partner organizations across leadership, management, and staff working with the data

San Francisco Needs for Analytic Tools and Data Reporting

In this section we provide an overview of data analytics and the recommendations about how San Francisco should build a system that will meet the information needs of stakeholders.

As OECE implements technical solutions for integrating administrative data, another important consideration is the *analytic capacity* to organize and learn from data. This will require intentional efforts to develop data warehouses that model data appropriately. Data modeling is the way in which data is stored and managed to enable access to the data for analytics. It involves extracting data from various systems, transforming it, and then storing the data in a manner that the data can be used and is useful, both dynamically and over time. Extracting and storing data in a useable format can be accomplished by building a *data warehouse*.

OECE is now preparing to procure new transactional systems which will provide a trove of data, but a central data warehouse with unified data model is needed to take advantage of this data. Data warehouses are separate from the systems to collect and manage production data. Program analytics and evaluation require data warehousing and modeling because both require access to data, *and also importantly* an understanding of the relationships between things such as enrollments, programs, workforce, families, and the context of services. OECE's and its partners' data goals must not only shape what information is being collected and why, but how it is stored and managed in a warehouse. In practice, this means that in addition to the data systems previously described that collect and manage information, OECE must also develop one (or more) warehouse(s) that define relationships among data elements and provide access.

Linking data modeling with program analytics is also important because it is not possible to anticipate all of the demands for data. Reporting and data requirements often change or evolve. Targets shift and are often redefined as priorities change and new information is brought to light. To align data solutions with this reality, OECE should devote time and resources to developing models that enable dynamic data analysis. The hardware, software, and labor requirements for data modeling must be carefully coordinated with OECE's IT support from the Human Services Agency.

Dynamic access to information (such as through ad hoc queries and data mining) will over time enable OECE to support effective planning, ongoing implementation, and rigorous analytics and evaluation of San Francisco's programs and services. Conversely, the absence of this function will create serious challenges and potentially even threaten the sustainability of OECE data systems.

We recommend the following next steps for developing analytic capacity to organize and report data. Currently data analysis is performed by OECE staff using STATA and data files that are requested from partner agencies. This process will be improved as the new system vendor begins to provide regular data extracts in standard file formats. Although still dependent on manual effort to transform and merge data files, OECE will be able to conduct internal analyses that support program needs and the learning agenda.

A long-term strategy for data analysis and reporting should include a data warehouse where all system data are modeled and stored. A robust data warehouse will allow OECE to develop many internal and outward facing analytic tools over time. In addition to standard reports, OECE should consider developing an ad hoc report tool so that data can be queried at any time as needed. Although at first, the data warehouse may only offer a limited number of reports, prioritization should be managed by the data governance structure once established.

Summary of Phases

OECE has a tight, but achievable, timeline to move forward on this project. To be successful in meeting OECE's long-term objectives, the efforts described above represent separate but coordinated functions that are needed to establish a culture of continual learning.

	Short-term	Medium-term	Long-term
	0-6 months (2018)	6-18 Months (2020)	19 and beyond (2021-)
Technical (collection and integration)	Phase 1: Family Matching	Phase 2: Program and Child Data Integration	Phase 3: Quality and Workforce Data Integration
Organizational (data governance)	Identify the necessary roles that will be included in the data governance structure	Identify the necessary policies and procedures that will need to be governed by the committees	Established data governance
Analytic (warehouse and BI)	Develop models	Build data warehouse	Implement Business Intelligence Tools

This roadmap provides context for San Francisco's work on early childhood integrated data systems with specific recommendations for closing the gaps in data availability and use. San Francisco, like many other cities and states, now needs effective strategies for positioning data and evidence as the centerpiece of policy and decision-making. This will require OECE leaders to first identify gaps in data availability to target efforts leading to effective and sustainable data use that ultimately will drive program and child outcomes.

San Francisco Data Governance Charter

OECE and Partner Vision Statement

The OECE data system will be a well-coordinated, integrated, and efficient system that allows families, programs, OECE, and its partners to use high quality, real-time data to promote favorable outcomes for children, age birth to five. The goal is not only to build an integrated data system, but to use it to inform decisions that support children and families in San Francisco.

Policy Statement

<This section should acknowledge data as a critical asset of the state and affirms the organization's commitment to implement data governance processes and structures. Defines the early childhood agencies' responsibilities to manage and use data in support of their mission and strategic plan.>

The OECE data governance committees are authorized to establish and enforce data governance standards and processes in furtherance of this goal. Establishing a set of data policies and procedures to ensure the responsible management and use of integrated data is a shared responsibility of OECE, Children's Council, Wu Yee, San Francisco USD, First 5. Data is an asset and the importance of accurate early care and education data for operations and reporting must be emphasized from the highest level of the organizations and policies and procedures must be adopted, communicated and enforced at all levels. The data governance standards and processes will reflect alignment with the Citywide ECE Plan and a learning agenda outlining the information needs across OECE and partner organizations.

Agency and Partnership Scope

<The purpose of this section is to describe the universe of data governed by the charter and organization encompassed by the charter.>

The data governed by the OECE integrated data system will be from the following partner organizations – Children's Council, First 5, OECE, San Francisco USD, Wu Yee, to address the learning agenda set forth by the executive leadership. The scope aligns to the planned integration of all data related to San Francisco's investment in early care and education, including data in the following data systems:

- Cocoa
- Early Learning SF (and the data migrated from the SF3C system)
- NDS shared data system

- WELS
- Wu Yee's Care Control 3 used for subsidy administration of OECE funding
- Children's Council's Care Control 3 used for subsidy administration of OECE funding
- Wu Yee's KidKare/Minute Menu system
- Children's Council's KidKare/Minute Menu system
- CA ECE Workforce Registry
- SFUSD Kindergarten Readiness Indicator data
- Other relevant data collection tools used by OECE and its systems partners

The key activities governed by the charter are as follows:

1. Development and implementation of executive, data management, and data steward committees.
2. Ensure all procedures developed under this charter and data governance manual are followed.
3. Ensure that the data analytics and reports created with the data are for or on behalf of the parties to improve agency services and analysis and reports will be shared with the respective parties prior to publication.
4. Ensure that no data will be released or reported in any manner that will allow the identification of any individual person, child, family, teacher, or site/provider, school.
5. To store data in a secure manner.
6. Maintain confidentiality of all educational records as required under the Family Rights and Privacy Act (FERPA), 20 U.S.C. 1232(g).
7. Maintain confidentiality of all health records as required under the Health Insurance Portability and Accountability Act of 1996 (HIPPA), Public Law 104-191.

Definitions of Key Terms

<The purpose of this section is to outline commonly used terms related to data governance (e.g., data collection, data request, data element).>

1. *Data* - the electronic representation of discrete facts.
2. *Data collection* – the process of gathering and measuring information on variables of interest in an established ,systematic fashion.
3. *Data dictionary* – a repository that contains comprehensive information about data.
4. *Data element* – a unit of data for which the definition, identification, representation , and permissible values are specified by means of a set of attributes.
5. *Data governance* – “Data governance is both an organizational process and a structure. It establishes responsibility for data, organizing program area staff to collaboratively and continuously improve data quality through the systematic creation

and enforcement of policies, roles, responsibilities, and procedures. There can be program data governance (single data system) or integrated data governance (multiple data system contributors)." U.S. Department of Education, National Forum Guide on Education Statistics (2011)

Governance Roles and Responsibilities

<This section should briefly describe the primary roles (individual and group) that will be established or formalized as part of the data governance effort, and the key responsibilities of each. Note: the responsibilities of each role can be further fleshed out in the Early Childhood Data Governance Manual.>

The data governance structure is outlined in Table 1 and described below:

The Executive Leadership Committee. This is a group responsible for developing policies related to the data infrastructure and how data will be collected and used. This group should include executive leadership across system agencies partnering with OECE to provide services to children and families. The executive director of OECE referenced an existing group of leaders from Wu Yee and Children's Council and that others including San Francisco Unified School District and First 5 San Francisco should be included in meetings going forward. This entity was established to discuss key issues related to the system-level services delivered by Children's Council and Wu Yee. With the membership expanded as described above, data governance key issues will be added to their scope biannually, with a goal of moving toward quarterly meetings.

The executive leadership committee is responsible for establishing and sustaining data governance. It does this by:

- Designing the learning agenda for the ECE system in San Francisco
- providing input on the data governance manual;
- creating the data management committee and appointing committee members;
- resolving issues escalated by the data management committee;
- providing input on data policies and major data-related decisions proposed by the data management committee; and
- holding all partner organizations accountable for participating in the data governance program and adhering to the data governance charter and manual.

Individual executive leadership committee members are responsible for:

- modeling and advocating data governance principles across OECE and partner organizations;

- communicating with the data management committee members who report to them to prepare for each executive leadership committee discussion and decision;
- including data governance responsibilities in the performance expectations of their staff who are members of the data management committee and data stewards committee; and
- holding all their staff responsible for adhering to the data policies and processes established through data governance.

The Data Management Committee. This committee will be the most important for OECE and the partner organizations. The data management committee is responsible for making decisions regarding policy implementation and operations for specific functional areas and proposing system-wide policy and operational approaches to fulfill the learning agenda. Members in individual organizations from Children's Council, First 5, OECE, San Francisco USD, Quality Partners, and Wu Yee are responsible for supporting integrated data system applications that create, capture and update data elements, and incorporate, edit and validate checks to protect data integrity. The data management committee makes policy recommendations to the executive leadership committee. Overall, the committee is focused on establishing and maintaining policies and processes regarding data collection, quality, availability, and use. The data management committee does this by:

- establishing, documenting, and enforcing data standards, policies, and processes (e.g., data request policy and process);
- identifying, prioritizing, and supporting the resolution of critical data issues that affect the quality, availability, and use of data;
- establishing and monitoring the progress of workgroups of data stewards to research and propose recommendations for addressing prioritized data issues;
- escalating to the executive leadership committee controversial, high visibility, or high-risk issues for resolution and policy recommendations for approval; and
- holding their program area staff accountable for participating in the data governance program and adhering to the data governance charter and manual

The data management committee is chaired by the data governance coordinator, and is made up of organization managers who are responsible for program areas that require the collection, management, and reporting of data, as well as those with data-related responsibilities.

Individual data management committee members are responsible for:

- modeling and advocating data governance principles with internal staff and other external stakeholders of their program area;

- attending data management committee meetings and participating in the decision-making process;
- ensuring that their program area data stewards and other staff adhere to the policies and decisions of the data management committee;
- sharing with the data management committee data issues, regulations, plans, and policies from their program area that may have an impact on other programs; and
- communicating with their organization staff the activities and decisions of the data management committee.

Data Stewards. Given the staff and partners arrangements in SF, the data stewards may not need to meet as a group, but they should be consulted by the data management/policy committee representatives on data elements and technical questions prior to changes being implemented. The data stewards are those that know the data and the quality of each piece of information. The knowledge they have is invaluable in the success of designing and using an integrated data system. Although they may not meet, they should have a clearly defined role and understand the responsibility they have to inform the data management/policy committee.

Data Stewards are responsible for:

- monitoring how data are defined, collected, quality assured, and reported to meet the requirements and the agency's and external stakeholders' data use needs;
 - informing calculation of metrics and determination of the source system of record for external reporting
 - reviewing and approving data releases
- communicating organizational needs for data privacy, security and archiving;
- documenting core processes regarding data collection, calculation, and reporting;
- ensuring metadata is documented and maintained;
- identifying critical data issues that impede data quality and use and participating in work groups to identify, propose, and implement resolutions;
- tracking federal, state, and other pending legislation or regulations involving data elements in his/her domain and communicating the potential impact;
- regularly evaluating and proposing process changes to improve data quality and/or agency efficiency;
- supporting their data management committee member by communicating data governance policies, processes and decisions to others within the program area;
- contributing information and expertise to support issues and discussions in data governance meetings.

Data Governance Coordinator. The data governance coordinator is a critical role to the implementation of data governance. In order for various decisions to both escalate and be

implemented (as shown in figure 1) there needs to be an individual who coordinates and communicates across all committees. The Data Governance Coordinator has an OECE-wide perspective and understands how data use should support and inform OECE's policies and programs and the larger learning agenda. This could be a new role, built within a current role, or while the position is vacant an external partner could fill the data governance coordinator position. In addition, the coordinator:

- facilitates the escalation and resolution of issues among the data governance groups;
- serves as a member of the data management committee;
- convenes working groups of data stewards to address critical data issues that span multiple organizations.

The knowledge, skills, and experience needed to be successful in the role of Data Governance Coordinator include:

- an understanding of how data use should support and inform education policies and programs, and an understanding of IT concepts and systems (but does not require an IT background);
- the ability to balance OECE-wide goals and strategy with making and implementing specific decisions in alignment with the broader goals and strategy;
- project management knowledge and meeting facilitation experience;
- experience leading process development and revision;
- training or mentoring experience;
- strong interpersonal skills, including the ability to mediate multiple perspectives and come to a resolution;
- strong communication skills, especially in translating between program areas and technology;
- understanding of data quality principles and strategies; and
- solution oriented.

Governance Activities and Processes

<Briefly describe the core processes that will be part of the early childhood data governance effort Note: The detailed steps of each of these processes should be defined and documented in the data governance manual.>

The operations of the data governance effort are further delineated in the Data Governance Manual, which will be developed as a companion to this charter. The Manual will cover, at a minimum, the following eight key areas, and can be expanded:

1. Scope –What are the most important questions of the learning agenda? What data will be included/integrated?
2. Integration – How will data be brought together and normalized?
3. Refresh – How often and how will data be updated to inform the learning agenda?
4. Changes – What is the input process for changes needed to inform the learning agenda? What is the input process for changes to data standards and systems?
5. Access – Who will have what level of access to the data to inform the learning agenda?
6. Confidentiality- How will data be kept confidential and secure while using data to inform learning?
7. Requests – How will external requests for data be reviewed and responded to?
8. Releases – How will reports on insights based on the learning agenda be disseminated?

Signatures

This charter remains in effect until June 30, 2019, at which time the Executive Committee will review the charter. The charter will be reviewed annually in June from that point forward. The charter establishing data governance will undergo annual review each June by the Executive Committee. The charter remains in effect as long as it serves the needs of the data governance effort and appropriate revisions can be made to support the functioning of the charter.

By signing this charter, the executive leadership is authorizing key personnel as outlined in the charter to establish data governance across OECE partner agencies to support the design and implementation of an integrated data system.

Director, Office of Early Care and Education

Executive Director, Children's Council

Executive Director, Wu Yee

Early Education Director, SFUSD

Executive Director, First 5 San Francisco