



To: Todd Bland, California Department of Social Services
Renee Mollow, California Department of Social Services
Raquel Givon, California Department of Social Services
Frank Mecca, County Welfare Directors Association of California
Christiana Smith, County Welfare Directors Association of California
John Boule, CalACES
Thomas J Hartman, CalACES
Diane Alexander, CalWIN
Michael Wilkening, California Health and Human Services Agency

From: The Alliance to Transform CalFresh, the Health Consumer Alliance, and the National Health Law Program.

Date: January 25, 2019

Re: Recommendations for Optimizing Data Management in the SAWS Migration

We appreciate the opportunity to engage in the SAWS migration planning efforts through the CalSAWS Stakeholder Engagement Meetings. In addition to our participation through this process, we have identified several areas of particular importance to our coalition of advocates (see the [SAWS Migration Interest Areas and Planning Priorities](#) for more information), and are looking forward to sharing our ideas in the months to come.

The attached white paper is a discussion of data-related challenges in the existing SAWS systems that have an impact on consumers. The paper includes recommendations for addressing those challenges in the context of the SAWS migration. Some recommendations are focused on operational capacity and protocols within the SAWS and ancillary systems, and others identify opportunities for the California Health and Human Services Agency to provide leadership to support data-driven work directly related to the SAWS migration context. These aspects of data management are equally important, though they may fall under the scopes of different departments.

We understand that many of the challenges identified in this report are known issues to the state and to counties, and thus migration planning may already include ways to address them.

Along with these suggestions, we request to learn more from your departments, agencies, and workgroups about any existing plans to improve data quality and analysis capacity to drive program improvements, and to have an opportunity to continue to participate in that discussion on an ongoing basis.

We look forward to your feedback on this white paper and learning more about the process in our upcoming SAWS Stakeholder meetings.

Optimizing Data Management in California State Automated Welfare Systems to Drive Program Improvement

January 2019



Acknowledgements

This report would not have been possible without the generosity of many people willing to share their time and expertise. We are grateful to those that helped us to understand the current landscape and future opportunities so that we could shape these findings and develop strong recommendations to improve California's capacity to leverage data about the CalFresh program to drive program improvements.

Becky Gershon, California Association of Food Banks
Brian Kaiser, California Department of Social Services
Dan Kelly, San Francisco Human Services Agency
Jonathan Koch, San Francisco Marin Food Bank Volunteer
Vakil Kuner, Office of the Chief Information Officer
Bridgette Lery, San Francisco Human Services Agency
Claudia Page, Consultant
Tracey Patterson, California Food Policy Advocates
Hilary Seligman, UC San Francisco
Peri Weisberg, San Francisco Human Services Agency
Mark Woo, Consultant

About the Alliance to Transform CalFresh

The Alliance to Transform CalFresh is composed of leading non-profit organizations that have come together to advance an ambitious but pragmatic goal: move California from a 66% participation rate in CalFresh (among the last in the nation) to an 80% participation rate, with no county below 70% by the end on 2019 in order to significantly reduce hunger in our state. Founding members include the California Association of Food Banks, California Family Resource Association, California Food Policy Advocates, Los Angeles Regional Food Bank, SF-Marine Food Bank, Western Center on Law and Poverty, and Catholic Charities of California.

Principal author: Diana Jensen, San Francisco Marin Food Bank, with research support from Claudia Page, Consultant.

For more information, visit TransformCalFresh.org or contact Diana Jensen at djensen@sfmfoodbank.org.

Table of Contents

Executive Summary	4
Background	9
Consumer Impact of Data Management Weaknesses	10
Building a Data Framework that Supports Continuous Improvement	11
Current Challenges and Recommendations for Collecting the Proper Data	12
Current Challenges and Recommendations for Data Accessibility: Extracting and Transforming Data into Usable Formats	15
Current Challenges and Recommendations for Using Analytics to Move from Data to Insights	19
Conclusion	21

Executive Summary

California provides health and human services benefits to more than 13 million Californians using infrastructure managed at the level of the counties. Counties use one of three software systems, called Statewide Automated Welfare Systems (SAWS), to handle enrollment and eligibility determinations for several core programs, including CalFresh, CalWORKS, and Medi-Cal. California is the only state to maintain multiple systems for this purpose. The state has recently begun the complex process of migrating these three SAWS into one centralized system, referred to as “CalSAWS”. CalSAWS is intended to support the provision of CalFresh, CalWORKS, Medi-Cal and other public benefits by 2023. This multi-year undertaking, which will cost more than a billion dollars, offers an unprecedented opportunity to examine the strengths and limitations of the current SAWS in order to optimize the new system to best serve California’s most vulnerable individuals and households.

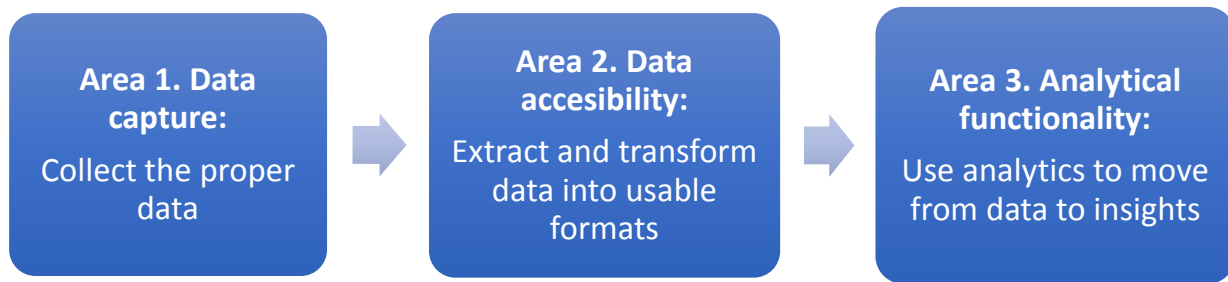
While there are many moving parts to this migration initiative, this paper tackles the importance of managing data collection, data access, and data analysis in driving continuous program improvement, especially as it relates to improving benefit program participation and consumer experience.

Given the types of data that are captured on both the individual and household level in the SAWS, it is reasonable to expect that counties and the state should be able to leverage the data to improve program participation, efficiency, and customer experience. Strong data systems should:

- Support county and state program staff in identifying opportunities for increased efficiencies for both clients and eligibility workers,
- Allow for analysis of the effectiveness and uniformity of implementation of policies and procedures across and within counties, and
- Identify inequities in outcomes across populations (e.g., by geography and sub-groups).

In truth, the current SAWS data management systems at the state and county levels lack the consistent data quality management, data access structure, and data analysis functionalities needed to support the goals of continuous program improvement. It is critical for the state and its partners to prioritize the management, access, and data analysis functionalities to drive program improvements for both administrators and program clients. These functionalities should be a part of the state’s SAWS planning, procurement, transition, and ongoing maintenance activities.

This paper lays out three focus areas for improving data management:



For each of these three areas, this report provides concrete examples of existing challenges at the state and local levels in the context of CalFresh, and recommendations for resolving them. Recommendations include those related to specific areas of the CalSAWS procurement that should be reviewed and strengthened to ensure strong data management functionality, as well as recommendations for the California Health and Human Services Agency (CHHS) to provide clarity on key performance indicators and the mechanisms necessary for all stakeholders to track them and understand their drivers on an ongoing basis.

CHHS' own [CHHS Data Playbook](#),¹ including the [CHHS Master Data Management Strategy](#),² as well as the agency's efforts related to the CHHS Open Data Portal, demonstrate the Agency's commitment to developing systems that enable data-driven continuous improvement. Those documents include a variety of great resources for implementing many of these recommendations.

Our recommendations are as follows, with further detail provided in the body of this report:

Recommendations for improving data collection and consistency:

Recommendation 1. CHHS should leverage the lessons from the [CHHS Data Playbook](#) to clearly define goals & objectives for CalFresh and other CalSAWS programs, identify data needed to assess progress toward them, and ensure adequate resources to maintain strong data in CalFresh and all other benefits cases. Specifically, HHS should:

- 1. A. Convene stakeholders to gain consensus on key performance metrics;
- 1. B. Develop a data governance framework to establish effective oversight of all areas of data quality management; and
- 1. C. Invest in state-level staffing focused on maintaining data quality and consistency across counties.

¹ California Health and Human Services Agency. *Data Playbook, Edition 3: A Department's Toolkit to Using Data*. Available online: <https://chhsdata.github.io/dataplaybook/>

² California Health and Human Services Agency (April 28, 2018). *CHHS Master Data Management Strategy*. Available online: <https://chhsdata.github.io/dataplaybook/documents/CHHS-Master-Data-Management-Strategy.pdf>

Recommendation 2. The Office of Systems Integration (OSI) should ensure that the CalSAWS procurement includes clear requirements regarding:

- 2. A. Expectations for future uses of data captured in the system;
- 2. B. Practices for developing data capture mechanisms;
- 2. C. The role of the CalSAWS project in data dictionary development;
- 2. D. Maintenance and related training to ensure data consistency; and
- 2. E. Expectations for the evolution of needs related to key data elements that will need to be accessible to users.

Recommendations for improving data accessibility:

In line with the guidance provided in CHHS' [Data Playbook](#) and its [Master Data Management Strategy](#), we recommend that CalSAWS procurement and contract development incorporate the following:

Recommendation 3. Codify the California Department of Social Services' (CDSS) and the California Department of Health Care Services' (CDHCS) access to client- and case-level SAWS data for analytic purposes, with agreements in place to govern details such as the scope of access, frequency of data sharing, training of staff as needed, and guidance regarding publication of findings.

Recommendation 4. Include an expectation that “canned reports” may need refinements based on local practice and testing before finalization. These should also provide the state and counties with annotated query language and consortia-level availability to respond to questions.

Recommendation 5. Requirements related to the creation of and access to data extracts and data tables, including:

- 5. A. Requirements for APIs or other mechanisms that allow for regular data extracts and live data tables to be retrievable in easily-useable formats that allow for meaningful analysis of key program functions. Review of the content and quality of these standard data extracts should be a regular process that includes county and state stakeholders to ensure they meet key management and evaluation objectives.
- 5. B. Providing a regularly updated reporting database that allows for ad hoc exploration of relevant variables.
- 5. C. Requirements for regular review and improvement of standard data extracts in consult with county and state stakeholders to ensure they meet key management and evaluation objectives.
- 5. D. A user-interface should be available for conducting data queries in order to empower additional users who may not have experience querying relational databases.
- 5. E. CalSAWS data tables should be developed anticipating the need to link SAWS data to other data sources (e.g. ancillary systems, customer satisfaction surveys, other state databases).

- 5. F. Access to the “live” comprehensive data tables should be the default, and not incur an added fee for counties.
- 5. G. The state might also consider purchasing data warehousing services through a separate procurement process, independent from the systems implementation vendor.

Recommendation 6. With regard to documentation, the CalSAWS procurement and contract development should include stipulations that:

- 6. A. CalSAWS should provide regularly-updated, user-friendly documentation linking the worker interface to the underlying data tables and data dictionary, including details about which elements are included in regular data extracts vs. “live” tables.
- 6. B. Data dictionaries should be developed and maintained at statewide system level, with an efficient mechanism for identifying and addressing concerns about local discrepancies.
- 6. C. Access to data dictionary information should be designed with the intention of making it as easy as possible for staff to see definitions in real-time when needed, and for new analysts to understand the meaning and expected usage of these variables and ensure consistency in analysis and interpretation across counties.

Recommendations for improving data analysis functionality:

Recommendation 7. CDSS and OSI should seek to ensure the following functionalities via the CalSAWS and/or related procurement processes:

- 7. A. Development and ongoing evolution of internal and public-facing dashboards that reflect key performance indicators as identified by CDSS and core stakeholders.
- 7. B. Built-in ability to drill down on key sub-categories (populations such as elderly/disabled, ABAWD, families, etc. as well as application sources, language and race/ethnicity, presence of earned income, geographic indicators, local office, etc).
- 7. C. Ability to leverage the unified CalSAWS to analyze data across public benefits programs.
- 7. D. Easy connections of CalSAWS data sources to a variety of business intelligence products to facilitate development of new analyses by a variety of users.
- 7. E. Standardized introductory and ongoing training for state and local analysts on using CalSAWS data with common analytical software, data dictionary guidance, and opportunities to connect analysts to CalSAWS experts and one another for peer learning.
- 7. F. Mechanisms to ensure that smaller counties have reasonable access to data analysis resources (e.g. via regional sharing of staff resources, access to state level staff, or otherwise).

Note that while these functionalities could be achieved as a part of the core CalSAWS system procurement, it may be worth considering the value of more modular procurements for analytical functionalities (e.g. similar to the role that the California Child Welfare Indicators Project provides for the state’s Child Welfare systems). To that end:

Recommendation 8. CDSS and OSI should assess alternatives for best achieving the data analysis functionalities above within the context of the CalSAWS procurement, other potential procurements or contractual arrangements with research institutions, and CDSS internal organizational capacity.

Background

California provides health and human services benefits to more than 13 million Californians using infrastructure at the level of the counties. Counties use one of three software products, called Statewide Automated Welfare Systems (SAWS), to handle enrollment and eligibility determination. The state has recently begun the complex process of migrating these three SAWS into one centralized system, referred to as “CalSAWS”. CalSAWS is intended to support the provision of CalFresh, CalWORKS, Medi-Cal and other public benefits by 2023. This multi-year undertaking, which will cost more than a billion dollars, offers an unprecedented opportunity to examine the strengths and limitations of current SAWS in order to optimize the new system to best serve California’s most vulnerable individuals and households.

California counties are currently organized into three separate management consortia, according to which SAWS IT system each county uses:

System	Counties	Approximate State Caseload
LEADER Replacement System (LRS)	Los Angeles County	30%
Welfare Client Data System (WCDS or CalWIN)	18 counties	40%
Consortium IV (C-IV)	39 counties	30%

Consolidation of these three systems into a single CalSAWS system will occur over the next five years, with LRS converting to CalSAWS first. C-IV will then migrate to CalSAWS, followed by CalWIN.

The pool of stakeholders in the SAWS migration initiative is broad. Counties have independent governance and operational structures, under which they manage SAWS procurements, hold contracts, manage vendors, and direct day-to-day system operations. The California Welfare Directors Association (CWDA), a nonprofit association representing the human service directors from each of California’s 58 counties, advocates on the behalf of county human services directors and their management of the SAWS systems. The California Health and Human Services Agency (CHHS), California Department of Social Services (CDSS), California Department of Health Care Services (CDHCS), and California Office of Systems Integration (OSI) each play a role in project management and oversight. Federal agencies, including USDA’s Food and Nutrition Services (FNS) and the Centers for Medicare and Medicaid Services (CMS) also provide federal oversight regarding certain aspects.

While there are many moving parts to this migration initiative, this paper tackles the importance of managing data collection, data access, and data analysis in driving continuous program improvement, especially as it relates to improving CalFresh program participation and consumer experience. The Alliance to Transform CalFresh (ATC) recognizes the enormous challenges in administering programs for millions of diverse Californians and appreciates the opportunity to offer insights. Because of the ATC’s focus on CalFresh, most observations in this

paper are related to that program, but the concerns and associated recommendations are relevant across all human services programs in the SAWS system.

Recommendations and findings for this paper are derived from discussions with state and local stakeholders and a review of existing reports, meeting minutes and presentations. It is not intended to be an exhaustive review of data challenges and opportunities, but rather an illustration of concrete examples of current CalFresh data challenges and the resulting limitations to enacting data-driven improvements. Given the complexity of existing planning efforts, it is likely that some of these recommendations may already be under consideration.

Consumer Impact of Data Management Weaknesses

California has one of the lowest SNAP participation rates in the nation, ranking 45th among the states;³ CDSS estimates that 72% of eligible individuals (2016 data) receive SNAP benefits.⁴ This means nearly three out of 10 Californians eligible for SNAP do not participate in the program, although rates vary substantially from county to county and some have participation rates as low as 41%. The state's SNAP participation rate is particularly poor when it comes to reaching two key subgroups of Californians: California ranks 50th out of all states at reaching eligible, low-income seniors,⁵ and 48th when it comes to reaching the working poor.⁶ Some of the variability in participation is due to the burden created for clients and eligibility workers by application and recertification processes. For example, many eligible applicants are denied benefits when they are unable to navigate the process; others choose not to apply at all because they perceive that it is "just not worth it." Language barriers or household circumstances such as the lack of a fixed address or use of an unreliable phone can exacerbate these experiences.

SAWS systems capture individual- and household-level data about program applicants, as well as measures of the application experience (such as time from application to determination, verification submission details, interviews, and reasons for approval or denial). Active cases include even more data, including benefit amount, notices, and reporting details. It's not enough to have the data; eligibility systems need to be procured and maintained with the expectation that the data they collect will be used to better understand program delivery and implementation outcomes.

³ Cunnyngham, Karen (January 2018). *Reaching Those in Need*. United States Department of Agriculture. Available online: <https://fns-prod.azureedge.net/sites/default/files/ops/Reaching2015.pdf>

⁴ California Department of Social Services Research Services Branch. *2016 County Program Reach Index*. Accessed December 12, 2018 at <http://www.cdss.ca.gov/Portals/9/DSSDB/Dashboards/CalFresh/PRI2016.pdf?ver=2018-07-23-092239-503>

⁵ National Council on Aging. *Senior SNAP Participation Visualization*. Accessed December 12, 2018 at <https://www.ncoa.org/economic-security/benefits/visualizations/senior-snap-participation/>

⁶ Cunnyngham, Op Cit.

Given the types of data that are captured in the SAWS systems, it is reasonable to expect that county and state analysts should be able to leverage the data to improve participation, efficiency, and customer experience. A robust data system should:

- Support county and state program staff in identifying opportunities for increased efficiencies for both clients and eligibility workers,
- Allow for analysis of the effectiveness and uniformity of policy or procedure implementation across and within counties, and
- Identify inequities in outcomes across population (by geography and sub-groups).

The current SAWS data management systems at the state and county levels lack the consistent data management, access, and analysis functionalities needed to adequately support the goals of continuous program improvement. It is critical for the state and its partners to prioritize the management, access and data analysis functionalities to drive program improvements for both administrators and program clients. These needs should be a part of the state's SAWS planning, procurement, transition and ongoing maintenance activities.

Building a Data Framework that Supports Continuous Improvement

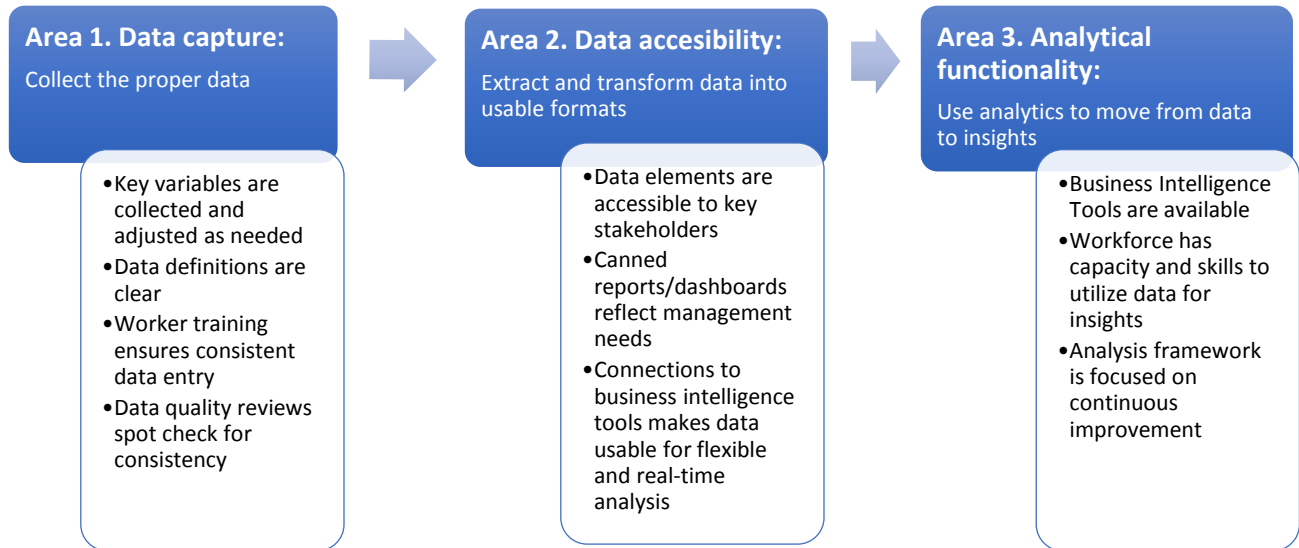
Migration to a unified CalSAWS system provides an opportunity to improve data management, access, and analytic capabilities with the goal of better supporting continuous program improvement. In this section, we articulate several key components of an effective data framework,⁷ provide examples of data challenges that are symptoms of underlying weaknesses in the current SAWS systems, and recommend solutions for addressing those data challenges in the context of the SAWS migration and related planning in order to drive continuous improvement.

What should a strong data management framework do?

Strong data systems and analysis reveal challenges in program access and operations, allowing for continuous improvements to *policies and procedures*. Timely access to data and useful analysis tools allow for effective data mining, including multiple levels of inquiry so that local and state administrators can drill down to specifics. Importantly, data systems and analysis can also reveal weaknesses in the underlying data systems (e.g. inconsistent data definitions/data entry, key data elements that would be useful but are not included in regular extracts, etc.),

⁷ Data management encompasses many more topics than are discussed in this paper. This paper highlights several areas that are especially critical for ensuring that the CalSAWS system can effectively leverage the data collected therein. One useful resource related to Data Management is the DAMA-DMBOK Functional Framework, available online here: <https://dama.org/sites/default/files/download/DAMA-DMBOK2-Framework-V2-20140317-FINAL.pdf> Additional resources are available in the CHHS Data Playbook, available online here: <https://chhsdata.github.io/dataplaybook/>

allowing for continuous improvement of the *data system itself*. A high-functioning data framework should support the outcomes illustrated below:



For each of these three areas, we provide examples of existing challenges in the context of CalFresh, and recommendations for resolving them.

Current Challenges and Recommendations for Collecting the Proper Data

Area 1. Collect the proper data

- Key variables are collected and adjusted as needed
- Data definitions are clear
- Worker training ensures consistent data entry
- Data quality reviews spot check for consistency

Each piece of data collected adds burden to program clients and workers. Therefore, data collection should be optimized to ensure that data can serve both program process purposes as well as analysis purposes. This requires:

- Clarity from the program team and other stakeholders on the data elements most likely to be useful for identifying program performance improvement opportunities;
- Documentation and oversight of precise definitions and examples of those data elements;
- Training to ensure that data is collected consistently regardless of differences in business processes across jurisdictions; and
- Periodic quality control procedures.

In practice, existing SAWS systems appear to be lacking in several of these areas.

Examples of data collection and consistency problems

Data that would allow for evaluation of several key parts of the CalFresh application and/or recertification processes are either not captured, or are captured in ways that are difficult to analyze across counties or consortia. This makes it difficult to conduct analysis that could lead to improvements for clients or efficiencies for counties in those areas of program administration. See sidebar for more specific examples.

Inconsistent collection of certain key variables may also make it difficult for California to easily comply with FNS' Major Change Rule,⁸ which requires that states provide regular reports to FNS on the impact of any "major changes" they make in their operation of SNAP (CalFresh).

Even within existing consortia where database fields are the same, weak data definition documentation or usage, differences in training, or differences in business processes limit the ability to compare data across consortia and counties. Always wondering whether differences in metrics are the result of true program differences or differences in data definitions, it is nearly impossible for the CDSS or counties to use data confidently to identify problems.

Recommendations for improving data collection and consistency

1. **Recommendation 1.** The CHHS should leverage the lessons from the [CHHS Data Playbook](#) to clearly define goals & objectives for CalFresh and other CalSAWS programs, identify data needed to assess progress toward them, and ensure adequate resources to maintain strong data. Specifically, HHS should:

A few specific examples...

- Denial and discontinuance reasons have historically been measured, coded, and reported differently across the three consortia for CalFresh, making it difficult to assess trends or opportunities. One of the biggest levers for increasing program participation is ensuring avoiding denials that process, not eligibility related. It is also a resource issue for counties: denied applications do not increase caseload, and therefore incur workload costs without additional resourcing.
- Differences in protocols for collecting variables related to language can make it difficult to analyze differences in consumer experience for English and non-English speakers.
- The CF 296 monthly CalFresh report summarizes application outcomes, including the share of applications denied due to procedural reasons. Several counties report virtually zero applications denied for procedural reasons, which suggests either a data quality issue or a business process that results in cases never being coded as procedurally denied.

⁸ <https://www.law.cornell.edu/cfr/text/7/272.15>. Examples of major changes include replacement of the state's automated systems, such as that planned with the CalSAWS data migration, as well as operational changes that are likely to continue in California such adjustment to call center technology, online portals, IVR systems and the like. Impact analysis include payment and negative error rate, program access, timeliness, customer service outcomes, and impact on certain types of households.

- 1. A. *Convene stakeholders to gain consensus on key performance metrics* for CalFresh and each of the other major public benefit programs to be managed by CalSAWS, both in terms of program compliance and excellence of customer experience. Stakeholders should include state and local subject matter experts as well as SNAP policy researchers. These performance metrics would help to identify a core set of CalSAWS and ancillary system variables that would be necessary to evaluate program performance, policy implementation, and/or impact of major operational changes on CalFresh access, equity, and ease of use.⁹
 - 1. B. *Develop a data governance framework* to establish effective oversight of all areas of data quality management, including consideration of data captured in ancillary systems,¹⁰ to guarantee improved data quality, coordination, more accurate reporting, and opportunities for research that can drive program improvement.
 - 1. C. *Invest in state-level staffing* focused on maintaining data quality and consistency across counties in the CalSAWS system to ensure that this foundational work is prioritized and maintained over time. Investments in staffing for counterparts at the local level may also be necessary.
2. **Recommendation 2.** The state and counties have the chance to hold vendors accountable for clearly defined data management requirements and vendors should be required to demonstrate capacity on this front as part of the procurement process. OSI should ensure that the CalSAWS procurement includes clear requirements regarding:
- 2. A. The expectation that data captured in this system must be able to be used for evaluating program implementation, assessing the impact and outcomes of county or state-level operational pilots, and the like.
 - 2. B. The use of data capture systems that minimize opportunities for error and inconsistency especially for key variables, as guided by CDSS. These may include pull down menus, error messages, and range checks.
 - 2. C. The role of the CalSAWS project in development and maintenance of strong data dictionary resources.
 - 2. D. The role of the CalSAWS project in the development and ongoing implementation of staff training resources and data quality checks to ensure ongoing consistency in data collection.
 - 2. E. The expectation that key data elements necessary for evaluating program outcomes are likely to change over time in step with program policy changes and evolving understandings of the possible drivers of program success.

⁹ See Appendix A for examples of the types of variables that are likely to be of interest

¹⁰ When captured consistently, data from ancillary systems related to topics such as lobby management, workflow, staffing details, or environmental factors can be extremely valuable in assessing program performance. Because significant variation exists across counties with respect to usage of these systems, discussions of related data governance will need to take into account those nuances.

Current Challenges and Recommendations for Data Accessibility: Extracting and Transforming Data into Usable Formats

Area 2. Extract and transform data into usable formats

- Data elements are accessible to key stakeholders
- Canned reports/dashboards reflect management needs
- Connections to business intelligence tools makes data usable for flexible and real-time analysis

For data to be useful to drive continual program improvement for both customers and county staff, the right data elements must be easily extractable in usable formats to the appropriate people in a timely fashion. Some stakeholders will benefit most from thoughtfully developed, pre-designed reports or dashboards, while others will need an easy way to extract data on an ad hoc basis.¹¹

In practice, the existing SAWS systems have often fallen short of these data accessibility objectives.

Examples of state access challenges

CDSS has not historically had easy access to county or consortia data. CDSS receives pre-set monthly and annual aggregate reports (not client- or case-level data). In order to receive additional data, CDSS must submit a specific request to Consortia via SIRFRAs (SAWS Internal Request for Research & Analysis). Because this is a lengthy process, CDSS has not been able to analyze or respond to data in real time, significantly limiting their ability to assess trends, identify inconsistencies, and support quality improvement efforts. Furthermore, like the monthly and annual reports, SIRFRAs are also typically aggregate summaries, without accompanying client or case-level data. It is difficult (and in some cases impossible) to assess the impact of various CalFresh policy changes and pilot projects in the current environment.

Limiting state access to aggregate data has historically prevented CDSS from:

- Conducting secondary analyses
- Reviewing potential reasons for data inconsistencies

¹¹ It is worth noting here that this paper does not address questions of data privacy protections, as that is not the focus of this discussion. Of course, such protections are of utmost importance for the SAWS systems given that they contain myriad personal information about program participants, and the CalSAWS system should be developed according to the highest standards in this regard.

- Disaggregating outcomes by race, ethnicity, primary language spoken, immigration status, or other household characteristics (e.g. age, children, earned income, student status, etc.)
- Effectively running scenarios to estimate the impact of potential policy changes related to potential legislative changes
- Developing a longitudinal database with SAWS variables for CalFresh, CalWORKS, or Medi-Cal
- Connecting SAWS data to other client-level data sources that would support research and evaluation of the impact of programs like CalFresh on health or other outcomes.

This not only has negative implications for clients in need of assistance, but it also presents challenges for the state in complying with FNS' recently enacted "Major Change" policy. Identifying, analyzing and reporting on the impact of major changes can be best achieved with appropriate data access.

As of the writing of this memo, we understand that CDSS has recently acquired access to some client-level SAWS data files for analytic purposes, with agreements in place to govern publication of findings. The Alliance to Transform CalFresh applauds this step toward making data accessible and useful for program implementation and quality improvement. CDSS has provided few public details on the contents of the data files, frequency, and reporting restrictions within this recent SAWS data-sharing agreement with the counties.

Examples of local access challenges

For counties, existing SAWS systems typically provide a suite of data products, the value of which can depend greatly on the local availability of analysts who know how to use them.

Local, consortia-designed, pre-programmed data queries and associated dashboards, sometimes referred to as "canned reports," provide aggregate information across a variety of domains. However, lack of information about the underlying queries can leave counties questioning the validity of the resulting data. When they are unable to get enough information to feel convinced of report quality, counties with stronger analytical staff capacity sometimes develop their own local versions of these canned reports. Because different versions of these canned reports often feed into state monthly reports, this results in a lack of consistency in variables that are aggregated at the state level.

Counties also typically have access to a variety of data tables on a monthly or daily basis with case- and individual-level data for applicants and active and discontinued cases. These data tables provide only limited insight into the functioning of the local business process with respect to application or report processing. Depending on the county and its vendor arrangements, those data tables may not easily to link to related data from ancillary systems (e.g. call center, task management, documents management, records management, etc.) or other program enrollment datasets. In theory, counties can link from standard SAWS data tables to a comprehensive set of "live" data tables (e.g. the CIS system in CalWIN) to pull in

other variables, but lack of clear documentation that connects these hundreds of data tables to the system's user interface, along with weak or non-existent data dictionaries, limits the true accessibility of these types of data products to a small number of skilled data analysts. When one of these individual analysts leaves their job, a county can become suddenly "data blind." Furthermore, access to the more comprehensive "live" data tables sometimes incurs an additional fee to counties.

Recommendations for improving data accessibility

In line with the guidance provided in CHHS' [Data Playbook](#) and its [Master Data Management Strategy](#), we recommend that CalSAWS procurement and contract development incorporate the following elements related to state data access, canned reports, creation and access to data extracts and data tables, and documentation:

- ✓ **Recommendation 3.** CalSAWS procurement should codify CDSS' and CDHCS' access to client- and case-level SAWS data for analytic purposes, with agreements in place to govern details such as the scope of access, frequency of data sharing, training of staff as needed, and guidance regarding publication of findings. Specifically, the state should be able to observe individual-level and household-enrollment in and out of programs like CalFresh in real time - or as quickly as the counties are able to observe them. The scope of data access should include as many variables as possible that might allow for analysis of operational differences that influence efficiency as well as equity outcomes, including consideration of data that may be housed in ancillary systems.
- ✓ **Recommendation 4.** CalSAWS contracts should include an expectation that "canned reports" may need refinements based on local practice and testing before finalization. These should also provide the state and counties with annotated query language and consortia-level availability to respond to questions to (a) ensure that figures reflect the management need, and (b) to enable state or local entities to use query language as a starting point to conduct additional related analyses without having to recreate the queries from scratch.
- ✓ **Recommendation 5.** CalSAWS procurement and contract development should seek to ensure the following with respect to **creation and access to data extracts and data tables**:
 - 5. A. Through APIs or other mechanisms, the state and local counties should have easy access to regular data extracts in easily-useable formats that allow for meaningful analysis of key program functions in order to make data-informed, continuous improvements. Depending on the user, these could be daily, weekly, and/or monthly extracts.
 - 5. B. Provide a reporting database at granular levels (e.g., case, application) that removes personally identifiable data elements but preserves relevant variables in non-aggregate levels to allow for arbitrary analysis not requiring per-analysis extracts. Continue updating this de-identified reporting data with new variables as they become identified as useful for future analysis.

- 5. C. Review and improvement of standard data extracts should be a regular process that includes county and state stakeholders to ensure they meet key management and evaluation objectives. State policy staff will need to advise on any new data elements that might need to be regularly used in the near future.
 - 5. D. In addition to data extracts, CalSAWS should incorporate a user-interface for conducting data queries in order to empower additional users who may not have experience querying relational databases.
 - 5. E. CalSAWS data tables should be developed anticipating the need to link SAWS data to other data sources to (a) comply with the Major Change Rule when relevant data may be stored in ancillary systems or customer satisfaction surveys, (b) evaluate the effectiveness of policy and/or operational initiatives, and (c) analyze the impact of CalFresh utilization on other health or economic outcomes.
 - 5. F. Access to the “live” comprehensive data tables should be the default, and not incur an added fee for counties.
 - 5. G. The state might also consider purchasing data warehousing services through a separate procurement process, independent from the systems architecture vendor.
- ✓ **Recommendation 6:** With regard to **documentation**, the CalSAWS procurement and contract development should include stipulations that:
- 6. A. CalSAWS should provide regularly-updated, user-friendly documentation linking the worker interface to the underlying data tables and data dictionary, including details about which elements are included in regular data extracts vs. “live” tables.
 - 6. B. Data dictionaries should be developed and maintained at statewide system level, with an efficient mechanism for identifying and addressing concerns about local discrepancies.
 - 6. C. Access to data dictionary information should be designed with the intention of making it as easy as possible for staff to see definitions in real-time when needed, and for new analysts to understand the meaning and expected usage of these variables and ensure consistency in analysis and interpretation across counties.

Current Challenges and Recommendations for Using Analytics to Move from Data to Insights

Area 3. Using analytics to move from data to insights

- Business Intelligence Tools are available
- Workforce has capacity and skills to utilize data for insights
- Analysis framework is focused on continuous improvement

With the right systems and data management in place, analysis can be done in myriad ways through querying, reporting, dashboards, ad-hoc requests, canned reports, predictive analytics, big data analytics, operational intelligence, data discovery, and many more activities. Insights garnered from the data can then be translated to action, which can be objectively measured and assessed for efficacy.

In the current systems, data are not consistently analyzed to drive insights and outcomes. The existing SAWS systems have not been procured and maintained in a way that seeks to maximize the analytical value of the data they collect.

Examples of state-level analysis functionality challenges

Over the last 5 years, CDSS has made significant strides in consolidating aggregate county-level reports using data from the SAWS systems, as well as data from the state's MEDS system, into a CalFresh Data Dashboard¹² that tracks a variety of key performance indicators statewide and across counties. CHHS has recently launched a similar product, the California Health and Human Services Program [Dashboard](#), which provides high level summary data on multiple CHHS programs. These are great first steps, and signal CHHS' growing commitment to fostering a culture of data driven innovation.

However, CDSS' analytical capacity has been hindered by an historical lack of state access to individual and client level records, lack of statewide consensus on key performance indicators and a centralized plan to evaluate performance drivers, and a lack of SAWS system functionality intended to support state evaluation work.

For SAWS system programs, there has historically been no institutional framework for statewide and county-level analysis of SAWS program data. CDSS' own Child Welfare programs provide a useful counterpoint: data captured by counties in the CWS/CMS child welfare system is transformed into longitudinal datasets and analyzed by the California Child Welfare Indicators

¹² Available online at: <http://www.cdss.ca.gov/inforesources/Data-Portal/Research-and-Data/CalFresh-Data-Dashboard>

Project, a collaborative venture between the University of California at Berkeley (UCB) and the CDSS. The project’s website describes its role as providing “policymakers, child welfare workers, researchers, and the public with direct access to customizable information on California’s entire child welfare system.” This example demonstrates a model in which analytical functionality sits outside the core programmatic database, with a separate contract that can specify analytical objectives for specific audiences, along with data analysis products (e.g. dashboards tailored to internal and external audiences), stakeholder engagement requirements, and protocols for supporting additional research pursuits.

Examples of county-level analysis functionality challenges

Counties often have their own, local level summaries, dashboards, and ad hoc analysis tools based on SAWS and other ancillary data systems. An analytical structure that encourages cooperation between counties could reduce total costs for counties in doing analysis, while also allowing for cross county comparisons to help drive improvements.

Furthermore, there is inconsistent capacity to conduct analysis using SAWS data from county to county. Different counties have different resources, and the state could benefit from increasing their staffing dedicated to this work to support counties or regional partnerships between counties. The departure of one skilled staff member at a local county can leave that county’s CalFresh administration “data blind” for months until new staff are hired and trained, and there is no guarantee that the departing staff will have documented his or her institutional knowledge to ensure a rapid return to analytical capacity when new staff are hired.

There is no meaningful community of SAWS county data analysts. County analysts work in 58 county silos; there is no formal structure for them to share learnings, queries, or analysis templates. They cannot easily build on each other’s work or on state level analysis to collectively improve thinking on how best to evaluate and improve program delivery, or identify ways that counties could better serve California residents through shared services. An institutional structure to create cross-county CalFresh data analyst collaboration — for example a standing CDSS or CWDA workgroup — could support cross-county knowledge dissemination, and will be particularly relevant given a single SAWS data system across all counties.

Recommendations for improving data analysis functionalities

- ✓ ***Recommendation 7:*** We recommend that CDSS and OSI seek to ensure the following data analysis functionalities via the CalSAWS and/or related procurement processes:
 - 7. A. Development and ongoing evolution of internal and public-facing dashboards that reflect key performance indicators as identified by CDSS and core stakeholders. Different audiences (state, county managers, supervisors, consumer advocates) require different kinds of dashboards to maximize their use.
 - 7. B. Built-in ability to drill down on key sub-categories (populations such as elderly/disabled, ABAWD, families, etc. as well as application sources, language and race/ethnicity, presence of earned income, geographic indicators, local office, etc).

Dashboards should be built on data sources that, by default, include common variables for analysis to make next-step queries as easy as possible.

- 7. C. Ability to leverage the unified CalSAWS to analyze data across public benefits programs. For example: assessing opportunities to address whole-person need, facilitating more efficient annual eligibility recertification/redetermination, identifying areas of unmet need, etc.
- 7. D. Easy connections of CalSAWS data sources to a variety of business intelligence products to facilitate development of new analyses by a variety of users. It should be easy for skilled users to take existing dashboards and modify them without the need to recreate from scratch.
- 7. E. Standardized introductory and ongoing training for state and local analysts on using CalSAWS data with common analytical software, data dictionary guidance, and opportunities to connect analysts to CalSAWS experts and one another for peer learning.
- 7. F. Mechanisms to ensure that smaller counties have reasonable access to data analysis resources (e.g. via regional sharing of staff resources, access to state level staff, or otherwise).

While these functionalities could be achieved as a part of the core CalSAWS system procurement, it may be worth considering the value of more modular procurements for analytical functionalities such as those embodied in the California Child Welfare Indicators Project model. Such models can also support data transparency through public-facing dashboards and the like. To that end:

- ✓ **Recommendation 8.** CDSS and OSI should assess alternatives for best achieving the data analysis functionalities above within the context of the CalSAWS procurement, other potential procurements or contractual arrangements with research institutions, and CDSS internal organizational capacity.

Conclusion

Unfortunately, the difficulty of making data-informed program improvements using the functionality of California's existing SAWS systems has led us to a situation where we are not getting what we pay for as a state. We pay the highest administrative costs in the nation¹³ for SNAP administration, and in return we have the 46th worst performance in getting needed CalFresh assistance to eligible low-income residents. Additional expenditures of state funding are unlikely to result in performance improvements without significant changes to the SAWS model. Programmatic improvements should be made based on the best available data analysis to make the most of federal, state, and county resources. All stakeholders would benefit from

¹³ Office of the Inspector General, United States Department of Agriculture (September 2018). *SNAP Administrative Costs*. Available online: <https://www.usda.gov/oig/webdocs/27601-0003-22.pdf>

increased performance from CalFresh and other benefits programs – especially low-income Californians who need these services to keep their families healthy. California would also be able to lower the state’s high poverty rate¹⁴ and recoup billions more annually in economic activity.¹⁵ We can’t make significant gains without a data management framework that supports continuous improvement.

¹⁴ Danielson, Caroline and Monica Brady (March 2016). *Just the Facts: The CalFresh Food Assistance Program*. Public Policy Institute of California. Available online: https://www.ppic.org/content/pubs/jtf/JTF_FoodStampJTF.pdf

¹⁵ Call, Jared and Tia Shimada (December 2016). *Lost Dollars, Empty Plates*. California Food Policy Advocates. Available online: <https://cfpa.net/CalFresh/CFPAPublications/LDEP-FullReport-2016.pdf>

Appendix A: Examples of data element types that are potentially useful for understanding program performance

The following list is intended only as an example of the kinds of data elements that may be useful for analyzing program performance and driving performance improvements. Some of these data elements are captured in the SAWS systems, and others are captured in ancillary systems. Data availability and usefulness of various data elements depends on local operations and thus would vary by county. This list is not intended to be prescriptive, but only to provide concrete examples of the kinds of variables that CDSS and counties might consider as important for meaningful analysis, and thus structure data systems in order to be able conduct analysis as easily as possible on these areas of inquiry.

HH composition/special populations:

- Basic demographics: age, race, ethnicity, language spoken, language reading
- Address
- Citizen/immigrant status
- Presence of children, and ages
- Disability status & types, any requests for reasonable accommodations
- Student status and related exemptions
- ABAWD status and related exemptions
- Earned income and type for each
- Unearned income and type for each
- New client vs. existing client
- Experience with welfare systems (previous client, client of other similar systems)
- Etc.

Application processing:

- Source (online, phone, in person)
- Connection to assistor
- Initial date application received by county
- Date signature received (electronic, paper, e-sig over the phone)
- Date eligibility determined, determination outcome, and reasons for denial as appropriate
- Interview process metrics
 - o Phone, in-person
 - o Date informed of interview time
 - o Actual interview time and scheduled time
- Verifications
 - o E-verifications
 - o Applicant-provided
 - o Requested due to "questionable"
- Call center metrics
 - o Call volume
 - o Time to wait, including excessive wait times

- Calls abandoned
- Calls dropped
- Calls with busy signal
- Customer satisfaction
- Etc.

Lobby Management

- When used (day of week, time of day)
- If self-service kiosk, is it attended? Did client need assistance?
- Kiosk ADA capability used?
- Language used on kiosk
- Length of check in time (e.g. 20 seconds or 2 minutes)
- Reason for visit (if captured)
- Wait time until next step after check in (if system has workflow)

Staff Information (may be linked to lobby management info)

- Position (e.g. clerk, EW, social worker, translator, etc.)
- Language used by staff

Workflow data (if workflow is available and used)

- Total visit time
- Total wait time / face time
- Number of interactions / number of waits

Analytics / BI data

- How did client get to the kiosk / office
- Transit access metric / walkability metric
- Weather conditions / traffic condition metric
- Visit data close to holiday?