Integrated Management of Survey Modes

Jerome Wernimont, Ray Snowden

Westat, Inc.
1600 Research Blvd
Rockville, MD 20850

Introduction

In efforts to improve response rates, reduce costs, and take advantage of emerging new technology, it is becoming increasingly common for organizations to utilize multiple survey modes to address different needs. Some studies use multiple modes at the beginning of data collection, while others adopt a phased approach or incorporate additional modes in an adaptive design. These multimode approaches present unique challenges in automating case workflow control, providing common services and supporting unified reporting and status monitoring.

In order to be efficient, the management of the survey process requires integrated systems that support end-to-end case workflow management and data transformation with a high degree of flexibility and real-time control. At the same time, such a management capability must interface with a variety of heterogeneous applications and platforms and address the needs of a wide variety of stakeholders including study managers, methodologists, and operations staff. This paper explores these challenges, the system development issues, and the approach taken and applied to projects at Westat.

Background

Surveys often involve the use of multiple collection modes such as hardcopy forms, Computer-Aided Personal Interviews (CAPI), Computer-Aided Telephone Interviews (CATI), and web surveys in order to achieve required response rates and meet cost targets. Data collection activities also generate various artifacts, including completed surveys instruments, hardcopy forms, biological specimens, and environmental samples that must be managed through post-collection processing. At Westat, each of these major data collection and processing activities, referred to as production lines, are supported by a custom management application that provides controls, services, and reporting specific to the nature of the activity.

Project survey designs vary significantly in terms of the sample make-up, study protocol, and data requirements. As a result, each project requires considerable flexibility and customization in the selection, configuration, and behavior of the production lines they employ, the rules by which data collection activities are activated and controlled, and the management applications they may use to support survey operations.

The ability to quickly transfer control and move data between various survey processes is a characteristic of increasing importance. Examples of where agility is important to survey operations include modifying the frequency or timing of reminders for a web survey, quickly shifting a respondent from a web-based survey to a CATI interview, or rapidly detecting possible patterns of data falsification.

Controlling survey activities across multiple modes is also needed to ensure that the study is executed as designed and to completion, i.e., that all cases in the sample pass through the study workflow in the same manner, following the same set of rules, and that all processes are completed.

Where an automated cross-mode management capability does not exist, the routing of case, transfer and conversion of data between production processes, and tracking and reporting of work status is left to the project to manage through one-off processing with a mixture of manual and automated features. Coordination between study operations and systems staff is used to control the release of sample, movement of cases from one mode to another, and manage case status. Visibility of the overall process requires reviewing a combination of reports provided by different systems. This approach can be error-prone due to the lack of automated process control, enforced workflow integrity, and comprehensive tracking and monitoring.

Information Technology Needs

Each data collection mode, processing function, and supporting service (i.e., production line) is typically supported by a management application that controls the processing model for the production line, captures and maintains

management data, and provides monitoring and reporting about processing status. For example, CATI production line management employs call schedulers and telephony systems; CAPI management supports case assignment and data synchronization to interviewers. One of the technical challenges for a multi-mode study involves integrating the different production line management systems in a flexible and responsive manner. In many scenarios, the lead time available to develop an integrated and complex solution is minimal due to contractual timing, IRB approval, or OMB clearance. Other factors such as late developing and shifting requirements, staff availability, and coordination between different operational units add schedule and cost risks to a project.

Excluding the actual survey data that is to be collected, the management processes rely on a significant amount of information, such as:

- Sample Data
- Participant Information
- Contact Information
- Organization Information

In addition, meta-data that drives the processes must be part of each system, whether in configuration or embedded in custom software, include data such as:

- Protocol Definition
- Status Values
- Task Outcome Definitions
- Timing Information

Production line management systems often represent a heterogeneous collection of technology including custom software, open source applications, and commercial-off-the-shelf (COTS) products and may utilize both legacy and emerging technology. At Westat, these systems have been developed over time, incorporate improvements that result from considerable project and operational experience, and in many cases are deployed independently.

Integration of these various systems has historically been done on an ad-hoc basis when required to support a specific study design and requires integration at multiple levels including data semantics and format, workflow harmonization, and technical interoperability. Custom services such as email, texting, and incentive management have also been implemented individually for projects as needed. In order to create an end-to-end solution, other post-collection processing activities also require integration through automation or manual operational processes, such as:

- Data editing
- Specimen receipt and tracking
- Quality Control
- GIS coding and review
- Data delivery

To address this issue, Westat has developed a multi-mode management system called M3 that provides an integration framework, an end-to-end workflow control mechanism, and tools and reports to permit projects to quickly design and implement complex and integrated study management systems that can meet their specific needs. M3 has been developed based on the following general principles:

- M3 is a platform that facilitates the integration of management applications. This permits Westat to continue to utilize existing production line management systems and to replace old systems and integrate new management systems with minimal disruption, i.e., once an application has been integrated with M3, it is interoperable with any other application integrated with M3.
- M3 has been developed incrementally over time and in close association with actual projects. With each new use of M3 on a project, features required for the project are added or extended and built into the baseline application for use on future projects.

- M3 is built to provide projects a great deal of flexibility. As much as possible, M3 behavior is data driven or rules based. M3 allows projects to add data items quickly and easily. M3 is architected to permit projects to integrate new applications quickly.

Development of M3

To support the development of M3, Westat selected an initial study that would help provide a real example of the needs of a multi-mode study. A second large scale study (with a longer lead time) was also used in parallel to help foster generalizations as development proceeded.

The development approach of a multi-mode management system needed to recognize the scope and complexity that the problem presents. No single management application provided support for the full survey life cycle. Rather, numerous specialized and heterogeneous applications provided support for the specific type of activities to be managed.

The approach Westat applied recognized several key things:

- The new system would not replace existing management application functionality unless this was otherwise needed and fit within the overall M3 architecture;
- Project requirements and decisions will come at the last minute, no matter how well-planned;
- Projects range widely in size and budget, from the simple and routine, to the large and complex longitudinal studies; and
- In order to permit integration among technically heterogeneous and diverse systems, M3 had to support multiple mechanisms for inter-system communication.

In Westat's goal of "Improving Research through Technology" the resulting platform needed to support the improvement of study execution, resulting in better response rates and/or lowering costs, for example. Some ways to achieve this improvement include:

- Providing a superset of integrated study management applications and functions that can be deployed quickly by projects with a great degree of flexibility;
- Permitting projects to make changes to study design during operations efficiently and with minimal disruption;
- Deploying a right-sized platform for managing study operations; and
- Supporting just-in-time decisions occurring later in the planning process without jeopardizing schedule or quality.

The development team worked with other IT teams supporting various data collection applications, and developed global definitions for common concepts (e.g., subject, participant, dwelling unit), as well as new abstract concepts that would support a distributed work architecture (e.g., production line, unit of work).

The resulting data model is a superset of data required by the various production lines, allowing for a centralized data transformation function and supporting the central repository of data needed for cross-mode reporting and control. The M3 data model includes concepts required to handle complex study designs across collection modes. Concepts are explicitly modeled where appropriate, but the architecture allows for ad-hoc additional properties that can be "transported" to other systems. Not all studies or data collection modes use the same subset of concepts from the M3 model.

This common model also simplifies mapping data accurately for data exchange, and enables interoperability between heterogeneous applications. This capability was a key to the eventual success of the platform as any single system just has to be integrated with M3 and it will be capable of exchanging data with any other system that is plugged in.

With regard to system architecture, these concepts also help define the responsibilities of each component, isolating complexity in the proper tier. For example, a production line node or adaptor is defined as the integration point between M3 and an external system. These nodes provide the appropriate technical interface between M3 and the production line (e.g., web service calls, file-based imports/exports, database integration), help manage the bi-directional flow of data between M3 and external applications, and perform the required data transformation.

At the core of M3 is its general purpose task manager engine that can model survey activities and data flows between data collection modes, embedded services, and post-collection processing. Using task specifications, task rules, and other meta-data, the task manager orchestrates activity, collects outcomes, and otherwise executes the study protocol. Cases and tasks are routed to or pulled from production lines through the production line nodes or adaptors. Data is exchanged as needed to provide baseline data or pre-loads to production lines, as well as capture management and study data such as contact information, or participant information.

The system was designed with a high level of configurability through meta-data tables, runtime code injection techniques, and other conventions for extensibility. Some of the configurability of the task manager includes:

- Rules can be altered quickly to improve response rates, reduce cost, or for other purposes;
- Task specifications and outcomes can have explicit or relative timing of tasks;
- Protocol applied can be based on waves, releases, subject type, or other characteristics; and
- Complex or custom logic can be injected to support unusual requirements.

The base set of functionality proved to be very effective and served as the foundation moving forward. Using the core features, Westat started deploying and enhancing M3 project by project.

Westat Studies

Westat has now deployed its M3 system on more than twenty (20) projects that vary from very small to multi-year longitudinal studies with hundreds of thousands of participants. M3 has helped these projects at Westat meet many of their challenges. Some of the key benefits that accrue to projects that adopt M3 include:

- M3 eliminates the need for each study to "stitch" data collection systems together using custom implementations for each project;
- M3 provides a platform for agile study configuration with configurable rules-based engine provides flexibility;
- M3 provides an integrated solution with high-end features with minimal costs to projects without significant customization needs;
- M3 provides a central repository of case management data; and
- M3 is a proven solution that has been validated and tuned through use on multiple projects.

The architecture of M3 assumed the responsibility of many of the integration complexities, including data transformations, and resulted in the simplification of configuring and deploying the production line management systems. The M3 team, with experience from many different projects now, supports and assists projects in customizing existing interface modules or developing new modules, and uses the experiences and requirements to feed the development cycle.

As an example, the first project that utilized M3 included the following characteristics:

- Paper forms
- Email notifications
- Web based data capture
- CATI follow-up for non-response
- Multiple waves

For this study, the rules for routing cases through the various processing steps were somewhat complicated, and the status values used by the different systems were not compatible. The introduction of M3 eliminated the complexities at the edges, leaving the three modes of data collection uncomplicated by the needs of each other. The team developed appropriate integration points suitable for each system and M3 automatically handled status translation, timing, and all of the coordination. The deployment for this study illustrated the savings from the non-linear complicating factors that were avoided, from software implementation to team communications.

Another field study that was a first adopter of the M3 system was a large longitudinal study and included the following components:

- Custom study management system
- Multiple waves
- CAPI production using Westat's field system (BFOS)
- Incentive disbursement through debit cards
- Specimen tracking production line using Westat's BEST system
- Data editing post processing
- Quality control using Westat's CARI system
- Product ordering service for field interviewer supplies

This project was an example where M3 could not attempt to solve the sample management or other study management activities, but could provide many standard services that were largely independent of specific study nuances. The project developed and maintained its own study management system (SMS), which is used to record and maintain complex sample relationships and manage the longitudinal study protocol. The project SMS is integrated with M3, which permits the generation and injection of cases that represent key longitudinal study tasks. New M3 services that were developed for this project included managing disbursement of funds to debit cards for participant incentives, managing product ordering from the field interviewers, as well as moving data around to post-processing services. The separation of the responsibilities allowed the custom SMS to be simpler in its design and focus on project specific management requirements.

Almost every study that has used M3 has had its own unique needs that have driven new features or new ways to configure the system. Over the course of the many studies, the system has supported study design changes in various ways, including:

- Simultaneous deployment to CATI and Web with safeguards to avoid unnecessary contact;
- Altering incentive amount mid-study;
- Altering notification templates of email or text messages;
- Extending time window for data collection modes;
- Adding additional reminders to participant communications; and
- Testing different study protocols and methodologies on different sample sub-sets.

Other techniques that have proven to be helpful when configuring and deploying M3 include:

- Pull methods used for mailings and other physical-based activities has allowed operational flexibility with no impact to the system;
- Just-in-time generation of user accounts to avoid pre-determination of credentials;
- Dynamic sample loading to production lines to avoid unnecessary data creation; and
- Dynamic incentive management for increased security and reduced waste.

M3's central repository has permitted project staff to track activities and status throughout the study operation life cycle, and has facilitated the needs for adaptive survey design and managing multiple protocol paths simultaneously. The core set of services such as email notifications, text messaging, incentive payments, and authentication continue to expand and have provided high-end features to projects that may not have been able to afford to include them all based on schedule or cost.

M3 Evolution

M3 is now a large suite of components and services and includes integration with most general use production line management systems as well as key post-processing services (Figure 1). This baseline platform can be used as a starting point for most projects and can be customized as required. Examples of some integrated components and capabilities include:

- M3 is integrated with Westat's BFOS for CAPI management.
- M3 is integrated with the multiple CATI schedulers including Voxco and custom applications.

- M3 is integrated with Westat's Teleform-based hardcopy processing.
- M3 is integrated with a variety of survey instrumentation including Blaise, Verint, and custom web instruments, and can manage the launching of web survey tasks and status synchronization.
- M3 is integrated with a number of post-collection processing applications including:
  - BEST for Sample and Specimen management;
  - CARI coding application used to perform QA using recorded interviews;
  - Survey data editing system; and a
  - Geospatial data review system to analyze field interviewer activity.

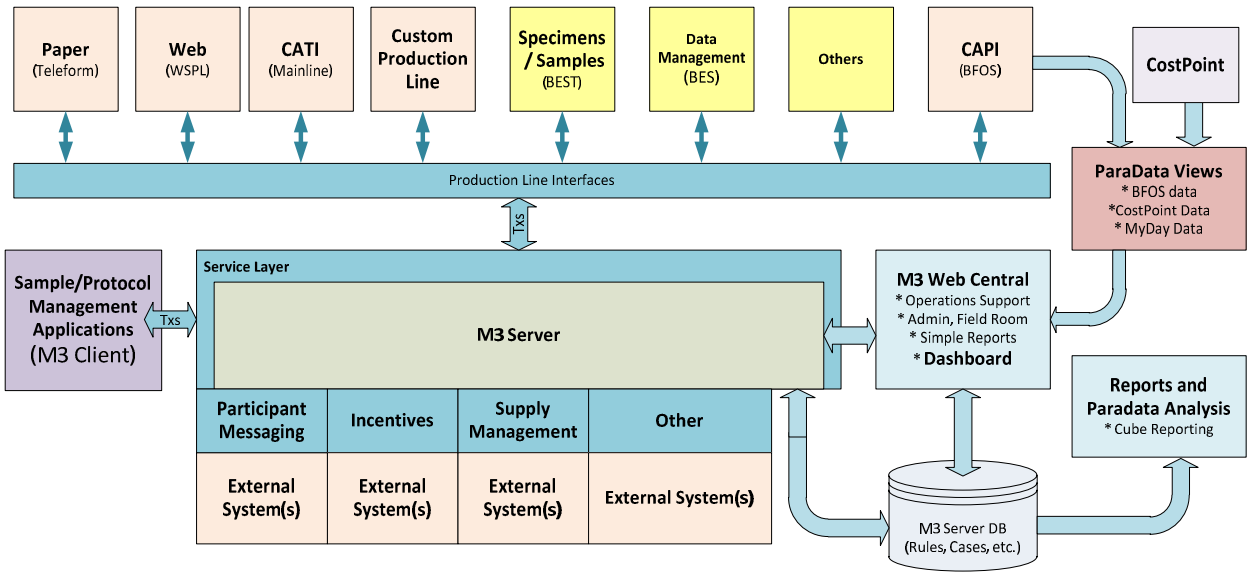

Figure 1: M3 System Platform Architecture

The incentive management service currently supports disbursements in the form of cash (through mailings), checks, debit cards, and Amazon gift cards that are dynamically acquired.

The architecture supports thin or thick study management systems (SMS) through the M3 client web service API. For large studies, the project SMS typically manages the survey sample and the study protocol. This ability is especially important for longitudinal studies and allocates the very detailed requirements close to the project, thus avoiding the over-reach and complexities of M3 solving every possible scenario.

The M3 web site has naturally evolved from the consolidated data model implementation and operational support features to include reusable and configurable web pages to support basic sample and study management. This natural evolution has enabled smaller projects to use M3 instead of manual management or developing their own system.

M3 is continuously validated, re-factored and incrementally extended as it is applied to each new project at Westat. Production line adaptors are developed to integrate new or custom project applications to M3.

Dashboards

The integrated M3 platform provided a foundation for the development of specific dashboards that support various operational needs. This is the next level of maturity for a multi-mode system. A common implementation of a dashboard includes graphical presentations of project status (snapshot) and historical trends of a project's key performance indicators. The content and organization of the data is designed to support the particular needs of a project. A well-designed dashboard will enable a user to understand the data at-a-glance and support quick and informed decisions.

At Westat we focused on an initial dashboard for study operations that addressed field operation needs, with quality, production, and anomaly tracking features. The dashboard utilized M3 central data as well as Para data developed over the years to support reporting and monitoring. Portlets were developed to utilize data and display in a concise graphical format, organized by field supervisor and their regions and staff.

State of the Systems Being Used

The current state of the multi-mode system has come a long way over the past few years and now can support reasonable levels of change in a quick manner, and is capable of reporting across collection modes. The capabilities need to continue to be expanded in many areas still, but that is to be expected as the potential for multi-mode continues to grow.

Analysis reports and other Para data reports should evolve into continual analysis and feedback. Visualization, managing changes, and other tools to help support adaptive design decision making should be explored. The ability for the project to visualize a protocol to ensure the correctness of its implementation before simulating case workflow through the system will speed deployment time.

Enabling project staff to alter study design without system staff intervention would provide more control closer to the decision makers. Introducing automatic alterations in workflow or mode based on sample priority or other characteristics would support embedded adaptive techniques. And for long running, longitudinal studies it is often a requirement to upgrade applications or adopt new applications or technology throughout the life cycle of the project.

M3 is architected and programmed to support the continuous evolution of technology and survey modes, and to be adept at supporting change and adaptive techniques. It has been built as a framework that facilitates the replacement or addition of systems with minimal disruption, as is often required for long running, longitudinal studies. The architecture supports expansion of its service offerings, and is able to integrate with heterogeneous applications through a wide variety of mechanisms. M3 has become the platform that provides both a full-functioned and integrated study management solution now and a mechanism for Westat to evolve and expand our capabilities for the future.