

# Integrating Clinical and Life Science Research

*Edgewater Technology provides custom development and systems integration to leverage data, processes and people across research and development initiatives.*

**Edgewater Technology's custom solution provides researchers with the ability to:**

- Customize the data collection and analysis environment for their clients' specific needs.
- Promote a culture of standardization and proactive adherence to best practices, yet enable a flexible methodology.
- Respond quickly and efficiently to new client demands and new project opportunities.
- Extend, refine and validate research teams' institutional knowledge base through the accumulation of active meta-data.

## *The Challenge of Clinical Data*

Edgewater Technology's client is a leading provider of clinical research reviews, data management and meta-analytic data and services to the healthcare industry. As an innovator in evidence-based medical information and an AHRQ (Agency for Healthcare Research and Quality) designated evidence-based practice center, this clinical research organization provides quantitative synthesis and qualitative assessment of clinical and pre-clinical data to support clinical trials program and protocol development, preparation of regulatory submissions and key clinical inputs to pharmaco-economic analyses.

To provide the highest quality service, the clinicians and biostatisticians in this organization staff must collect and assimilate large volumes of research and healthcare related data from a wide range of both publicly available and proprietary sources. A typical analysis might draw from dozens of pre-clinical summaries, patient care databases and clinical trials reports, as well as thousands of articles in the medical research literature. Their analysts must find a way to balance the need for strict adherence to rigorous scientific and statistical standards that ensure the validity of their findings, with variability in the quality and form of their data sources and frequent changes in the needs and research objectives of their clients.

## *Coping with Data Diversity*

Ask anyone working in life science or clinical research what the most pressing problem they face today is, and you will get the same answer time and again - too much data, available from too many sources, in too many formats, with too many possible meanings. The critical challenge faced by these organizations is two-fold. First, researchers must be able to

locate and obtain the data they need in a usable format. This is essentially a mechanical problem of mapping disparate data structures and formats between sources and destinations, usually manipulated in the form of computer systems, files, databases and applications. Second, they must be able to determine unambiguously the precise meaning of each data element obtained from a given source, eliminating any chance for misinterpretation or misuse of the information. This is fundamentally a semantic problem of capturing the exact nuance of meaning for each piece of data and preserving and enforcing that definition wherever that information is used.

Successful data integration is only possible when both the mechanical and semantic aspects of the problem are addressed. Coping with these issues while reviewing or monitoring thousands or even tens of thousands of information sources, in an environment with virtually no data standards, is a daunting task. Analysts needed the ability to establish and record the specific data collection requirements to support each client's particular research objectives. They also needed to capture the precise meaning of each data element to be collected, in a manner that would remove any ambiguity in its definition and ensure that its meaning would be stable and persistent over the life of the project and beyond. Furthermore, they wished to exploit their accumulating dictionary of data element definitions both within and between multiple projects. This would enable them to integrate their data collection and data analysis activities and continually improve their internal processes. It was apparent that a "one size fits all" solution was not going to work.

## *The Solution*

To cope with these challenges, researchers need a system and data architecture that is highly structured, but flexible to support the varied needs of the specific analysis they are asked to perform. They chose Edgewater Technology to create a custom solution offering a flexible design that can be tuned to the precise requirements of each new client protocol, while still providing a stable structure that supports the legitimate comparison and analysis of cross-project data. This feature is crucial when planning subsequent protocols that address the same or related disease settings, therapeutic agents or target outcomes.

The custom solution is based on a comprehensive design that includes a robust, active meta-data dictionary that facilitates researchers' needs to define, categorize and catalog the precise definitions of each required data element. Data definitions enforce standard names, data types, value constraints, units of measure, cardinalities and level of

aggregation for each data element. Relationships can be expressed between metrics recorded at various levels in the protocol, supporting direct drill-down from the overall project, to supporting detail at the summary, study, treatment group, treatment detail or patient level.

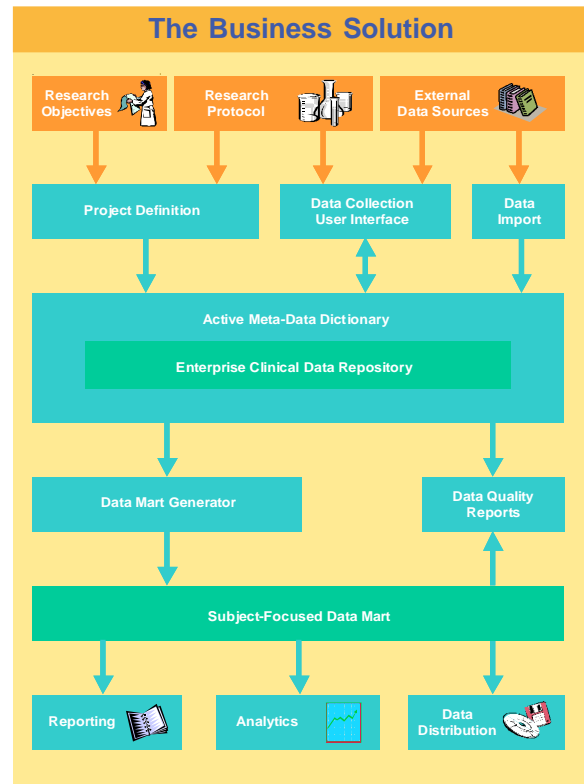
A project team of clinicians and statisticians reviews the protocol and the existing meta-data dictionary, and assesses the correspondence between the meaning of the data elements already defined and those required for the new project. Existing data element definitions that meet the precise requirements of the new protocol are reused, and legitimately new data definitions are created where necessary. These data management tools and the disciplined processes that these tools enforce combine to facilitate an ongoing standardization of shared nomenclature and data resources across the entire research organization.

Once the meta-data is recorded and activated, an integrated system of flexible application components goes to work for the users, tailoring and configuring every aspect of the system to the specific needs of the protocol. User interface screens are rendered using layout instructions that derive directly from the activated meta-data dictionary. Likewise, data collection screens, data extract generators, a protocol-specific reporting and analysis data mart, and other custom reporting, analysis and distribution tools are all tuned and driven directly on the basis of the activated meta-data. The result is a standard integrated repository that is automatically customized for each new research project.

## The Benefits

The integrated clinical data repository enables each research team to focus on their client's individual needs and customize the data collection and analysis environment accordingly. The combination of standard, structured data semantics and flexible solution components allows the scientists and clinicians to concentrate on their specific research objectives, without having to worry about the creation of a system and data infrastructure to support their needs. Using the active meta-data dictionary, researchers are able to tailor their system without intervention from technical support or engineering staff. This new system allows for maximum flexibility, while still promoting a culture of standardization and proactive adherence to best practices. Researchers are able to respond quickly and efficiently to new client demands and new project opportunities.

As their active meta-data accumulates, research teams are extending, refining and validating their institutional knowledge base, preserving this critical asset for exploitation and reuse on subsequent projects and protocols. Project teams are enhancing their ability to manage and combine data in meaningful ways, providing insights due to the integration of data. This facilitates the iterative and collaborative processes of data mining and analysis, and helps investigators identify



complex relationships and connections between facts reported separately from disparate sources. It also directly supports their need to perform statistical pooling of data from multiple trials and protocols, a hallmark of meta-analysis.

Project teams are able to mass customize their product and service offerings to meet the specific needs of their clients. By balancing standards with flexible delivery, research teams can plan new projects in a structured and disciplined manner, looking for opportunities for ongoing improvement and service excellence. They can build on their competitive advantage by expanding their customer service through tailored product offerings.

The new system has been leveraged on a number of research projects and client engagements, supporting systematic literature reviews, clinical data meta-analysis, and monitoring in a wide range of disease settings and therapeutic areas. It is an extensible system, supporting access to data through industry standard interfaces and formats, enabling researchers and statisticians to use various best-in-class analytical tools in their work. Additional capabilities can be added to the system in a straightforward manner, capitalizing on the meta-data services and interfaces for such features as secure and personalized Web portal access, XML data and meta-data distribution and Web services deployment.