

Data Warehouse Findings and Observations



Nexus Innovations, Inc.

State of North Dakota

State Court Data Warehouse

Version: 1

June 5, 2001

Prepared by: James Vollmer, Senior Consultant – Business Intelligence

james@nexusinnovations.com

EXECUTIVE SUMMARY	4
BACKGROUND.....	5
FINDINGS AND OBSERVATIONS.....	5
<i>Data Quality.....</i>	<i>5</i>
<i>Standard Reports.....</i>	<i>5</i>
<i>Capturing of Continuances.....</i>	<i>5</i>
<i>PCSS Integration.....</i>	<i>6</i>
<i>AS/400 Platform.....</i>	<i>6</i>
<i>Information Sharing.....</i>	<i>6</i>
<i>Court Calendars.....</i>	<i>6</i>
<i>Case Aging Reports.....</i>	<i>6</i>
<i>Profiling.....</i>	<i>6</i>
<i>Ad-Hoc Reporting - Statistical Analysis.....</i>	<i>7</i>
<i>PCSS Integration.....</i>	<i>7</i>
RECOMMENDATIONS	7
PROJECT OVERVIEW	8
PARTICIPANTS.....	8
ARCHITECTURE RECOMMENDATIONS	9
<i>Network Operating System - Microsoft Windows NT/ 2000.....</i>	<i>9</i>
<i>Data Warehouse DBMS- Microsoft SQLServer 7.0.....</i>	<i>9</i>
<i>Extract/Transform/Load Tools - DataMirror Transformation Server.....</i>	<i>10</i>
<i>Microsoft IIS Active Server Pages.....</i>	<i>10</i>
<i>Seagate Crystal Reports.....</i>	<i>11</i>
NEXT STEPS	13
MICROSOFT SQLSERVER 2000	13
SCHEDULE LOOKUP BY CASE	13
OBLIGATION LOOKUP BY CASE/INDIVIDUAL.....	13
AD-HOC REPORTING.....	14
NDCOURTS.COM INTEGRATION	14
FUTURE INITIATIVES	15
DIGITAL DASHBOARD	16
DIGITAL DASHBOARD ARCHITECTURE	16
<i>Operational Environment</i>	<i>16</i>
<i>Business Intelligence System.....</i>	<i>17</i>
<i>Digital Dashboard Services.....</i>	<i>17</i>
EXAMPLE DIGITAL DASHBOARD	17
FUTURE INITIATIVES CONCLUSION	18

EXECUTIVE SUMMARY

A goal of the State Court Technology Committee is to provide a single voice, strategy and point of contact for all court related technology issues. A critical component in meeting this goal has been identified as development of a data warehouse. To meet this need, Nexus Innovations has been engaged to evaluate and implement a data warehouse solution for the North Dakota State Court.

Data warehousing is not a new concept. However, it's still an organization's best chance to get valuable data for analysis quickly and reliably. Existing operational systems (and their databases) are not designed for end-user access and analysis, i.e., business intelligence (BI). By cleansing, simplifying, enriching and localizing data, end users get more value faster than operational systems can deliver.

Interviews were conducted with key individuals from the Supreme Court, the South Central and South West Judicial Districts, as well as court administrators to gain an understanding of desired deliverables from a data warehouse.

The interview process revealed several key areas that will be focused on as part of the first set of deliverables from the data warehouse: The areas are as follows:

- Enable easier data access by extraction, transformation and loading of key data files to a data warehouse platform.
- Enable easier ad-hoc reporting by utilization of enterprise reporting tools, which will enable the Court's IT staff to respond to both internal and external client needs.

The deliverables identified from the interview process are as follows:

- Development of a data dictionary and enterprise data model to reflect the contents of the data warehouse. This deliverable will aid both the Court's IT staff and clients in understanding the breadth and depth of data housed in the data warehouse.
- Evaluation and implementation of extract, transform and load (ETL) tools to move the data from the operational systems to the data warehouse.
- Development of court calendars, accessible via a web browser.
- Development of key statistical reports, accessible via a web browser.

Satisfying these deliverables will not only make the implementation of the data warehouse a success, but will provide a solid foundation for future development to facilitate information sharing for the Court.

Background

Today's criminal justice system is much different than twenty to thirty years ago. Because of changes in laws and litigation processes, today's judicial system needs to run in an efficient and effective manner. A parallel can be drawn between a judicial system and a supply chain system in manufacturing. To provide a high level of performance, all components and players need to know what each other are doing. Dissemination of information in a palatable manner is the key to management of this process.

The key to effective management of any supply chain is the capture of relevant data pertaining to the process. In the case of the State Court, this data is captured in the Unified Court Information System (UCIS), running on an IBM AS/400 housed in Bismarck. This system provides the administrative staff of the court access to all pertinent case and counts information, and also provides other administrative functions such as court scheduling and standard reporting. This system services approximately 80% of all cases in the state of North Dakota.

With this taken into consideration, several key users and sponsors of UCIS were interviewed to identify what data is required from an analytical and external aspect. This will provide input in regards to the recommendations from this document.

Findings and Observations

Data Quality

Both the IT staff and clients feel that UCIS performs an admirable job in regards to data quality. Examination of the data also revealed strong referential integrity (relationships) between the master and validation files, with the only issues being duplicate records and invalid dates identified in only a few data files.

? The court IT staff has not only corrected the data quality issues in UCIS, but has found the source where the records were created and modified the applications accordingly. The process and procedures which were developed to identify these anomalies will be re-used during the migration efforts of the Minot and Mandan municipalities, as well as migration of the Northeast Central Judicial district to the statewide UCIS system housed in Bismarck. The cleansed data will then flow into the data warehouse.

Standard Reports

Standard UCIS reports currently provide the core of analytical reporting for the North Dakota Court system. Although accurate, the reports lack flexibility in regards to format and delivery. Further, existing and new reporting applications need to be made available to a wider clientele. Moving these AS/400/RPG based applications to a web based solution would be difficult to accomplish.

? Adaptation and use of enterprise reporting tools that facilitate web publishing will enable flexibility in output formats (PDF, Excel) and enable the Court to deliver information to non-UCIS clients, such as abstracters, newspapers, and other interested parties.

Capturing of Continuances

Interviewees voiced their concerns on how continuances are currently captured in UCIS. This can have dramatic impacts on the court scheduling process.

? Although not addressed in this phase of the data warehouse, algorithms and processes can be developed to identify continuances and can be reported to administrators and judges using the adapted enterprise reporting tools. Analysis efforts need to be conducted to design these algorithms and processes.

PCSS Integration

All interviewees state that the PCSS (Cass County) integration needs to be addressed. Cass County accounts for approximately 14% to 20% of all cases in the state of North Dakota. From an analytical perspective, it would benefit the judiciary to systematically coalesce the information from these two different, yet similar systems. From an administrative perspective, integration of PCSS and UCIS is a must.

AS/400 Platform

The AS/400 platform is designed to provide on-line transaction processing (OLTP), and does so in fine fashion. However, it does not lend itself particularly well to the ad-hoc reporting. Further, data from an on-line transaction processing (OLTP) system is structured for high-performance on-line access. In most cases, this data structure does not lend itself to the ad-hoc querying required by the client community. While UCIS standard reporting may be adequate for day-to-day activities, it lacks the flexibility in formatting and output to be useful for other data analysis activities.

? The data warehouse will aid the Court IT staff in development and implementation of new analytical reporting by exploitation of features and functionality in the relational database systems and reporting tools adapted by the Court.

Information Sharing

Sharing of data between other state entities is becoming a major issue for the State Court administrative staff. At the present time, data is only exchanged between the Department of Transportation, State Health Department and UCIS. Upon review of the judiciary's IT plans, information exchange between state agencies will be a major issue in the next 4 years. The AS/400 platform does not lend itself to easy integration into heterogeneous environments.

Court Calendars

All interviewees identified the court calendar as an immediate need. Even though a calendar already exists in UCIS, the flexibility in selection and output does not lend itself to access to people other than UCIS clients.

? Although the calendar is a deliverable in this phase of the data warehouse, increased functionality and integration between widely accepted calendars tools will need to be addressed in future development efforts surrounding UCIS, its clients, and the data warehouse.

Case Aging Reports

Aging of active cases measures the effectiveness of a court system. An immediate analytical need of judges and court administrators is the ability to see how their court is performing in accordance to state and federal guidelines and the ability to identify cases that are not within these guidelines.

? An opportunity exists to measure the time between events in a court case, much like measuring time between events in a merchandising/distribution environment.

? Future development efforts may include the development of intelligent "agents" that identify cases prior to the case falling out of compliance. This information can then be delivered to administrators and judges for proactive action.

Profiling

With more and more cases being managed by UCIS, the ability to quickly identify a case that may be a problem will assist both judges and administrators in processing. This may be accomplished via analysis of the defendant, offense, and case events. As mentioned previously, profiling can become a powerful method to proactively manage issues relating to the court.

? As mentioned previously, development of profiling techniques can become a powerful tool which

will aid the Court in proactive management of cases.

Ad-Hoc Reporting - Statistical Analysis

Reporting needs not currently met by UCIS creates a backlog for the IT staff. This can be alleviated by adaptation of business intelligence tools that enable internal clients to "serve" themselves, thus providing pertinent data to assist in day-to-day operations of the Court.

In many instances, the data from the data warehouse becomes "married" with data from another system or application. These analytical efforts can help identify where the data warehouse needs to be extended to provide an accurate, concise data management system.

? Adaptation of enterprise reporting tools will enable the IT staff to prototype a report, have it verified by the client and moved to a productive environment in a drastically shorter timeframe than previously experienced.

? The data warehouse is an organic, living system. Similar to any organism, the data warehouse will grow over time if nurtured and fed. Although current data needs have been met, the increased need for other types of data will drive growth and development of the data warehouse.

PCSS Integration

All interviewees voiced their concern about PCSS not being included as part of the data warehouse effort. The warehouse may serve as the center for other development and integration efforts in the upcoming future, especially in regards to statewide criminal justice initiatives.

Recommendations

The recommendation is to progress with the development of a data warehouse, centered on development of court calendars and case analytics. The data used to populate the warehouse would be sourced from UCIS. The delivery mechanism for the reports will utilize off the shelf tools that will enable both IT and key clients to develop ad-hoc reports.

? The data warehouse will include an enterprise data model and data dictionary to assist knowledge workers in further development of reports and associated analytics for the court.

? Since PCSS and UCIS integration is being addressed separately, PCSS integration into the data warehouse will not be addressed at this time.

Project Overview

Participants

James Vollmer - Senior Consultant, Nexus Innovations

?? Provide overall project management and development of the data warehouse

Kurt Schmidt - Director of Technology, North Dakota Supreme Court

?? Executive Sponsor

Jerrold Arneson, North Dakota Supreme Court

?? Provide AS/400 and NT Server Technical Support

Kari Goos, North Dakota Supreme Court

?? Provide UCIS database knowledge and address data quality issues

Andrea Schmidt, North Dakota Supreme Court

?? Provide UCIS database knowledge and address data quality issues

Dale Zaspotuil, North Dakota Supreme Court

?? Provide AS/400 and NT Server Technical Support

Architecture Recommendations

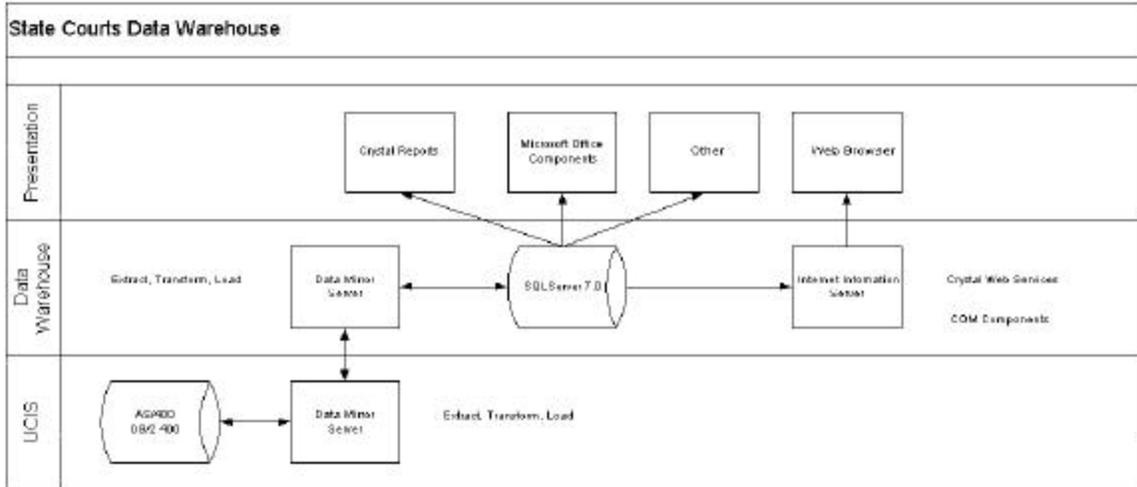


Figure 1: Architecture Overview

Network Operating System - Microsoft Windows NT/ 2000

Utilization of the court’s existing network operating system will not only leverage existing security, but will enable use of existing processes and procedures to administer application and database security. In addition, development tools available for the Windows platform facilitate rapid application development and integrate with the Court’s desktop offerings.

Microsoft’s Internet Information Server (IIS) will be used as a web application server. IIS is integrated into the NT environment, and thus existing user security profiles can be utilized in securing important components of the data warehouse and associated applications. Adaptation of this technology, however, mandates the use of Microsoft’s Internet Explorer as a browser.

Data Warehouse DBMS- Microsoft SQLServer 7.0

Today’s DBMS offerings give information technology directors many viable choices. With the Court’s adaptation of Windows NT as the network operating system, Microsoft SQLServer 7.0 is a logical choice for the data warehouse.

SQLServer provides a robust, high-performance platform for all database activities, specifically data warehousing and on-line analytical processing (OLAP). SQLServer’s integration with Windows NT Server and the Microsoft Office family makes integration into the State Court server environment a relatively simple process.

From a long term, strategic perspective, SQLServer is a core component of Microsoft’s .NET strategy, encompassing information exchange, collaboration, security, and application development. A key component of .NET is extensible markup language (XML), which is supported natively by the next version of SQLServer (SQLServer 2000). Although XML is less than two years old, it is quickly becoming a standard transport mechanism for data and application integration. Standards organizations are either developing or have developed XML-based schemas to facilitate data exchange between organizations.

Adaptation of SQLServer will enable the Court IT staff to leverage existing knowledge of Microsoft products and allow implementation of existing Windows NT security profiles to be integrated into the data warehouse.

Extract/Transform/Load Tools - DataMirror Transformation Server

DataMirror Transformation Server is a software solution that enables real-time transformation and flow of data. Implementation and testing within the UCIS environment verified the following claims stated by the vendor:

Peer-To-Peer Replication

Peer-to-peer architecture and native support for high performance TCP/IP communication protocols ensures direct database-to-database connectivity. No programming, data staging, or gateway technologies are required.

? This claim was verified via replication of selected UCIS data between Grand Forks County and a test Microsoft SQLServer.

Guaranteed data delivery

Asynchronous replication ensures guaranteed data delivery independent of two-phase commit logic. This prevents any possible transaction deadlock between multiple replication engines while providing the ability to recover data in the event of a communication failure. Unique real-time capabilities make data delivery a high performance reality.

? This claim was verified via examination of the files on the UCIS AS/400 and changes applied to the SQLServer database after the network connection was terminated between the two systems.

Change data capture

Whether data is replicated in real-time or on a periodic basis, only changed data is captured and transferred from the publisher to subscriber systems, eliminating redundant data transfer and improving operational efficiency.

? This claim was verified by selectively updating key data files between UCIS and the data warehouse.

Heterogeneous support

High performance, programming-free data integration between DB2, Oracle, Microsoft SQL Server and Sybase across Microsoft Windows NT/2000, UNIX, Linux, IBM OS/400 and OS/390. Replicate flat files on a refresh basis into relational database tables. Easily export non-relational structured data locked in legacy applications into client server and e-Business environments.

? As stated previously, UCIS data was replicated to 2 different platforms.

Built in transformation

Translate values, derive new calculated fields, join tables at either the source or the target for derived expressions, reformat field sizes, field names and data types, convert date fields and more.

? This process was verified via development of a date time field for the SQLServer database, which concatenated CCYYMMDD information stored in 4 different fields.

This software has a proven track record within other state agencies, including the Bank of North Dakota and the State Information Technology Department (ITD). Adaptation of this tool will not only assist in the development and deployment of this project, but will enable the court to share information with other state agencies, such as the Attorney General's office. An inexpensive purchase of software for the recipient's system will allow the court to replicate and/or share information from both the AS/400 and SQLServer platforms.

Microsoft IIS Active Server Pages

Microsoft's Active Server Pages are the server-side execution environment for IIS that enables the execution of scripts and ActiveX server components on a Windows NT server. By the combination of scripts and components, dynamic content and powerful Web-based applications can be easily developed.

ASP provides a very open development environment, with support for both Microsoft Visual Basic® Scripting Edition (VBScript) and JavaScript. Organizations can leverage the investments they already have in these scripting languages.

ASP ships as part of the IIS environment, free of charge. Development tools that can be used include Microsoft's Visual Studio, which is a GUI development tool used to develop Windows and ASP-based applications. Adaptation of Visual Studio will leverage the Court IT staffs' knowledge of Visual Basic for Applications (VBA), which is a subset of Visual Basic and used in productivity tools such as Excel, Word and Access. The skills learned by utilization of one language are easily transferable to the other.

Integration between IIS, SQLServer and the Microsoft Component Object Model (COM) also presents other possibilities due to their inherent interoperability. With the increasing acceptance of XML as a data interchange standard, the application server supporting creation of XML documents enables data exchange and opens up new and exciting possibilities in regards to flexibility of output types. The current release of IIS supports accessibility to the XML DOM (Document Object Model), allowing creation and formatting of XML documents via XSL (extensible style sheet language). In the case of this project, XML and XSL will be used to facilitate the creation of HTML (hypertext markup language). As products become available to facilitate the creation of other media types, the XML can be easily formatted to satisfy the requirements of the Court.

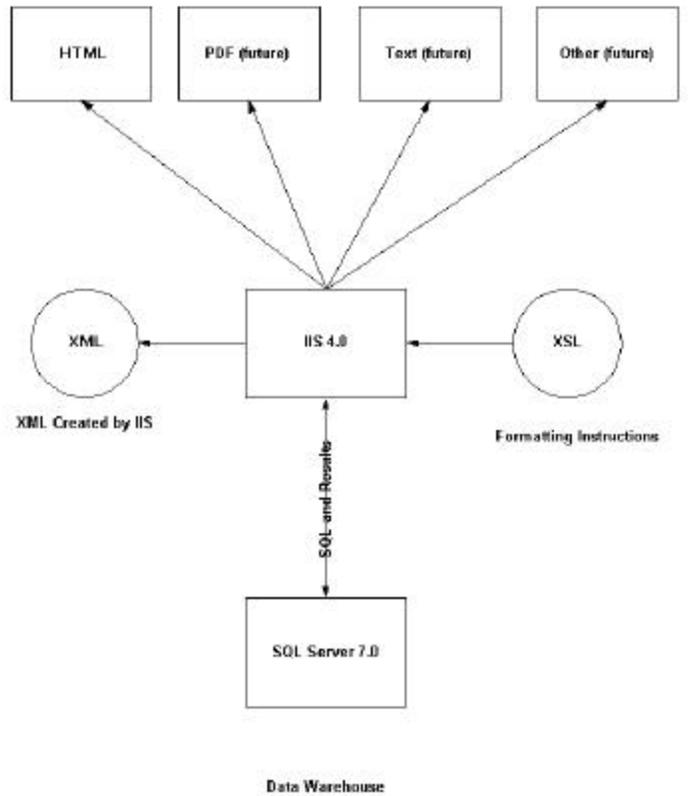


Figure 2: State Court Web Services Architecture

Seagate Crystal Reports

Seagate Crystal Reports is the most widely used report writer in the world, with an installed base of over 4 million copies. The software has a low acquisition cost, has flexibility in regards to output

and formatting, and is already in use the Court's IT staff. Crystal Reports also allows report developers to deploy output to the web, allowing clients to view this information using a web browser.

? Adaptation of this product set will not only leverage the Court's existing knowledge of the tool, but will facilitate information sharing by deployment of reports using a web browser.

Next Steps

The foundation is the most important architectural component of any construction project. With the implementation of the data warehouse and associated technologies, the state court has created a solid foundation on which to base future business intelligence and e-government activities upon.

Microsoft SQLServer 2000

Implementation of SQLServer 2000 will help the court achieve increased DMBS performance, and allow the adaptation of technologies such as OLAP and Data Mining that are integrated into the DBMS offering. SQLServer 2000 supports creation of XML natively within the database. This will assist the rapid development and deployment of web-based applications sourced from the data warehouse. In addition, SQLServer 2000 supports the use of XML as an update mechanism within the DBMS.

? These futures will not only enable the Court to stay with current technology, but will create a myriad of creative possibilities relating to the delivery of information.

Schedule Lookup by Case

Both external and internal clients have deemed delivery of Court schedules by case number a need.

? This process can utilize the court calendars developed as a deliverable from this project, with only slight modifications to inquire by case/count number.

Obligation Lookup by Case/Individual

Delivery of case obligations via electronic means has been deemed a necessity by the Department of Corrections. This mechanism will be used to identify whether or not an individual is meeting all obligations set forth by the court.

? This application can be created via leveraging the financial and condition information stored in the data warehouse.

Ad-Hoc Reporting

The true value of the data warehouse will be measured by the flexibility inherent in the structures contained within the database. Based upon the needs of a decision maker, flexible, prompt driven reporting can be created to meet their analytical needs. The tools chosen for report development can also be transitioned into the client community without extensive training, and utilization of an OLAP (on-line analytical processing) server and integration with spreadsheet can be utilized to access statistical information.

? An upcoming deliverable from the data warehouse should include a simple, prompt-driven, web-based reporting application to allow analysis of key attributes contained within the data warehouse. The analysis to be performed should include what these attributes are, and what prompts need to be made available to the client.

NDCourts.com Integration

In a recent address by Chief Justice Gerald W. VandeWalle to the Institute of Court Management, it was stated that the public is demanding a better system of access to justice. North Dakota is a leader in efforts allowing the public to access justice information. One of the keys to North Dakota's leadership has been Justice Dale Sandstrom's effort in the development of NDCourts.com.

During the interview process, Justice Sandstrom stated that he foresees the incorporation of data warehouse data into NDCourts.com. One of the first may be the viewing of judges' calendars from NDCourts.com, which could access the data warehouse. Since both NDCourts.com and the data warehouse use the same core technologies (Windows NT), the ability to access the calendar information would be straightforward. A suggested architecture is shown below.

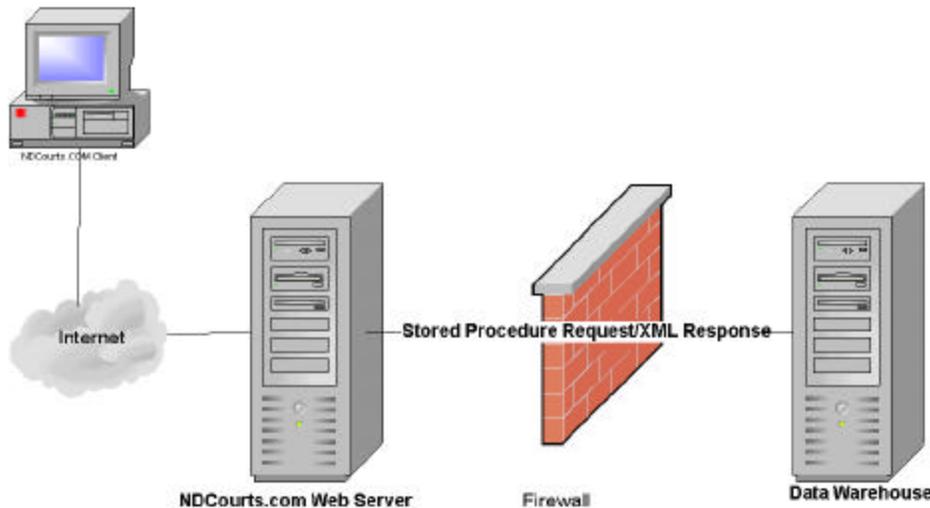


Figure 3: NDCourts.com / Data Warehouse Integration

Adaptation of this architecture will allow NDCourts.com real-time access to the calendars stored in UCIS. Further, the use of stored procedures and NT security will reduce the risk of malicious activity against the data warehouse and applications.

Future Initiatives

The data warehouse is one of many tools that assist in management decision-making processes (MDMP). The MDMP is a circular, cause-and-effect model consisting of just four steps.

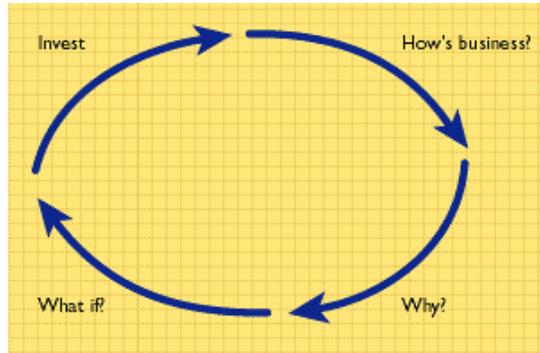


Figure 4: The Management Decision Making Process (MDMP)

1. **How's Business?** To show the state of the business, traffic signal highlighting (green, yellow, red), graphs, and charts can be used to show if key performance indicators (KPI) are good, bad, or on-target. These presentation tools use aggregate data from the data warehouse.
2. **Why?** Why is this indicator not within guidelines? This process involves delving into the details used to formulate the indicator and identifying what is wrong (or right). Terms such as "drilling into details" are often used in determining the "why" behind a business problem.
3. **What If?** Based upon data collected by the decision maker, different scenarios may be played out by performing analysis using spreadsheets and other analytical tools to model the effect of change on the process. This process usually involves data from other systems and processes.
4. **Invest.** Based upon analysis performed, the decision maker influences change onto the process.

The MDMP starts again to measure if the analysis and investments are paying dividends to the organization.

The MDMP is powerful tool by itself. However, the capture of non-structured information gives decision makers the ability to retroactively discover why a particular decision was made and gauge its impact (whether successful or not) on the process. This non-structured information includes emails, presentations, office documents and memos.

The support for collaborative processing and knowledge management is now becoming an integral part of business intelligence applications and has been coined **e-intelligence** by the industry. E-intelligence is primarily delivered to decision makers via a portal or digital dashboard.

Digital Dashboard

A digital dashboard is a personalized, customized solution for decision makers that consolidates personal, team, corporate, and external information. The digital dashboard provides single-click access to analytical and collaborative tools. It also brings an integrated view of an organization's knowledge sources to an individual's desktop, enabling better decision making by providing immediate access to key business information.

As with any industry terminology, the term digital dashboard has become overused and it's definition unclear. However, in the Court's view a digital dashboard could be viewed as a workbench, secured and individually tailored for the decision maker it serves. The following will address the individual architectural components, and how they relate to the Court.

Digital Dashboard Architecture

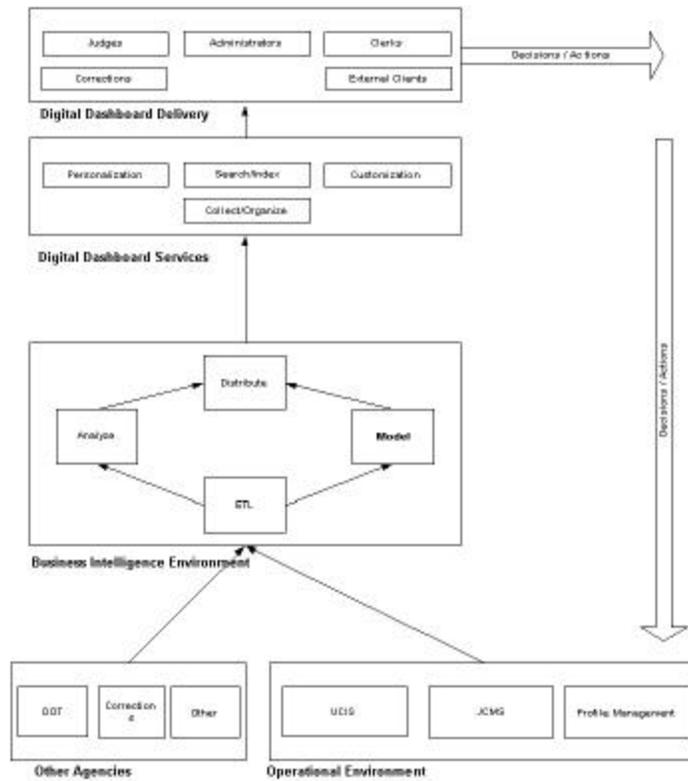


Figure 5: Digital Dashboard Architecture

Operational Environment

The operational environment(s) serve as the tool that is used by the decision maker to manipulate change onto the process. In the case of the state court, this system is UCIS. Opportunities exist for collaboration with systems within the Department of Transportation, Department of Corrections and JCMS. Collaboration tools such as personal information managers, electronic mail and contact listings also fall into this category. Future operational environment components could also include document management.

Profile management is an important component of any business intelligence system.

Profile management defines the decision maker from a systems perspective. Also, other methods of contact including e-mail, phone number are managed as part of the profile. In short, it defines the personalization characteristics of the decision maker.

Business Intelligence System

The business intelligence system is the foundation of the digital dashboard. It serves as a consolidation point for data pertaining to all decision-making. The majority the business intelligence system was constructed during the current phase of the data warehouse project. However, as with any data warehouse, the scope and breadth of data contained in the warehouse will continue to grow over time to accurately reflect the business.

Digital Dashboard Services

Digital dashboard services consist of 4 components:

1. *Collection and Organization.* The digital dashboard does not populate itself. Assigned administrators of the portal perform this duty. Administrators need not be IT personnel: in most cases, they are key players in the client community. Collection and population may also be automated process, such as scouring the business intelligence system looking for anomalies in the data.
2. *Security and Personalization.* Personalization is the key to any digital dashboard. Personalization becomes part of the user profile, which also includes the components that the decision maker subscribes to. In the Court's case, management of the profile would be integrated into existing directory services.
3. *Searching.* If you can't find what you are looking for, how can you make a decision? Searching capabilities are an integral part of any digital dashboard. More advanced search engines will allow the searching of non-structured data such as documents and presentations.
4. *Customization.* As the adaptation of the digital dashboard progresses, new functionality will no doubt need to be created. The digital dashboard needs to be easily extended to facilitate change.

Example Digital Dashboard

A digital dashboard is comprised of purchased or in-housed developed components called web parts. Web parts can include functionality such as e-mail and calendaring, analytics (reports, spreadsheets, charts) and other custom developed or purchased parts. To facilitate the MDMP, the delivery should include key performance indicators that decision makers would act upon, as well as personalization of the workspace to aid the decision maker in other daily activities.

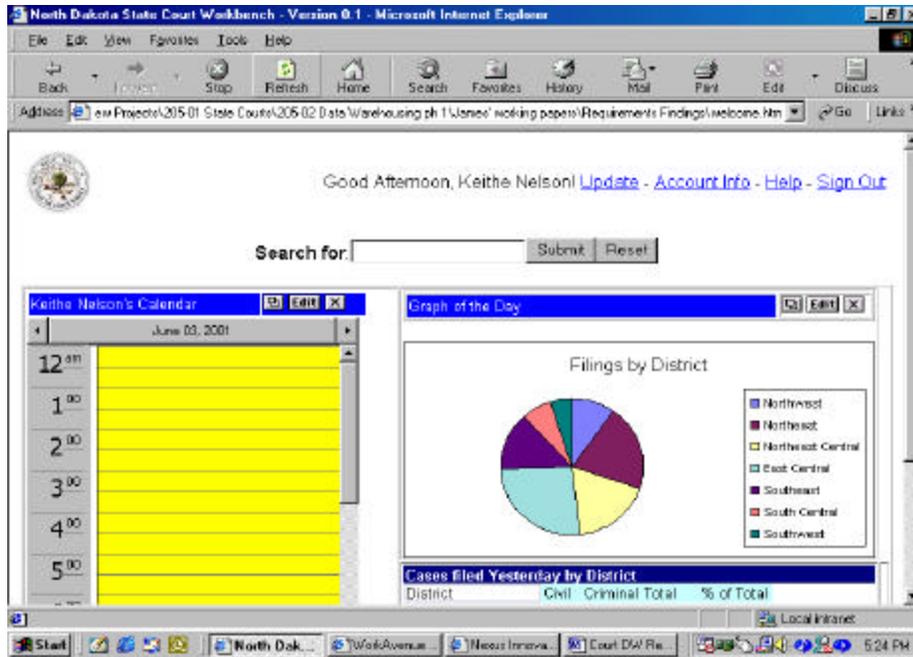


Figure 6: Example of Court Workbench

Future Initiatives Conclusion

A digital dashboard will become an important component in assisting all decision makers of the Court. The ability to examine a situation, discover what issues need to be addressed and implement change in a collaborative environment ensures that all interested parties are informed and can play a role in the decision making process.