WhitePaper

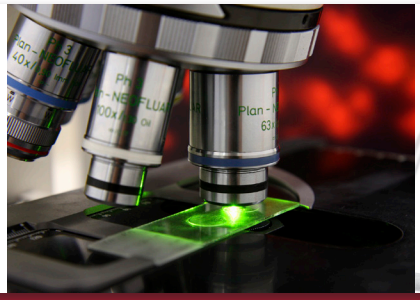
Closing The Medical Data Gap:

Using IT To Close The Gap Between Health Information Systems And External Documents

Author: Mark Terry

Editor: David Rasmussen






DARK Daily Laboratory and Pathology News @ darkdaily.com

Table of Contents

Introduction

This report will present an innovative technological solution to closing the gap between health information systems and external paper documents.

Healthcare creates a wealth of data. Every patient encounter, even a routine checkup, generates a significant volume of information ranging from personal identification data such as Social Security number, age, and address, to clinician notes and impressions, patient data like blood pressure, temperature, pulse, as well as complex laboratory information.

The trend is toward digital medical information, automated data entry, and the use of Electronic Health Records (EHR), Electronic Medical Records (EMR), Laboratory Information Systems (LIS) and Health Information Exchanges (HIE), among many other health IT tools. Healthcare reform and various government-led stimulus packages have further pushed the digital revolution onto healthcare. Nonetheless, there is a significant gap between reality and full adoption of digital medical information. A great deal of information is still produced on paper. Integrating it into an EHR for digital archiving, search, retrieval and analysis is a difficult, expensive and time-consuming task.

Although there are two broad types of medical information – structured medical data and unstructured medical data – this report will focus primarily on structured medical data. It will address laboratory health information, the complexity of structured laboratory data, and various approaches to integrating structured lab data into health IT systems along with their strengths and weaknesses. And finally, this report will present an innovative technological solution to closing the gap between health information systems and external paper documents.

Chapter 1:

Structured Data Versus Unstructured Data

Unstructured data is any data that does not reside in an organization's information systems.

Structured data (medical or otherwise) can be defined in several different ways, but for the purposes of this report, "structured data" is information that comes in numbers, tables and rows (or can be placed into numbers, tables and rows). Unstructured data, then, is generally found in the midst of a narrative, but can also include handwritten notes, charts and various medical images like x-rays and CT scans. You can further break down unstructured data into "textual unstructured data" and "non-textual unstructured data." Another way to define unstructured data is any data that does not reside in an organization's information systems.

Another aspect of structured data as discussed here is "external documents." An "external document" is any type of document or information that comes from outside the system's format. In other words, if the information comes via paper or needs to be manually handled or significantly modified to be used in the electronic system, it is an external document.

Raymond D. Aller, MD, FACMI, FHIMSS, FCAP, Clinical Professor and Director of Informatics in the University of Southern California's Department of Pathology and Contributing Editor for CAP Today, says, "Structured data is data that you can compute upon. Basically, in order to be structured, you need a computer to be able to look at the data and take unambiguous action."

Structured healthcare data preferably follows certain international guidelines and formats, including:

- SNOMED: Systematic Nomenclature of Medicine, which is a hierarchical medical classification system;
- IDC-9: International Statistical Classification of Diseases and Related Health Problems, 9th Revision;
- IDC-10: International Statistical Classification of Diseases and Related Health Problems, 10th Revision;
- HL7: Health Level Seven International, a global authority on standards for interoperability of health information technology with members in over 55 countries;
- First Data: An international drug database;

And others, including classifications set forth by the World Health Organization (WHO).

Unfortunately, not everything loosely classified as "structured data" has been structured in a useful fashion.

Unfortunately, not everything loosely classified as "structured data" has been structured in a useful fashion. Tom Murray and Laura Berberian, in a Computerworld blog from March 31, 2011 titled, "The importance of structured data elements in EHRs," say, "The general physician and hospital population have hastily 'automated' the paper process and in doing so, have compromised creating standards and charting using structured data. The end result is what exists today: patient medical records comprised of 'data dumps' from different systems, the seemingly insurmountable task of meeting meaningful use requirements set forth by the American Recovery and Reinvestment Act of 2009, and major obstacles in sharing healthcare information seamlessly between the entities within HIEs."

Chapter 2:

The Push For Digital Medical Data

"The HITECH Act was incorporated into ARRA to promote the adoption and meaningful use of health information technology."

On February 17, 2009, President Barack Obama signed into law the American Recovery and Reinvestment Act of 2009 (ARRA). Part of the spending authorized $20 billion to help develop health information technology (IT) infrastructure. Among the numerous things it did was formally codify the Office of the National Coordinator (ONC), endorsed certification and standards, and updated the Federal Health IT Strategic Plan. From 2011 to 2015, physicians are eligible for financial bonuses for demonstrating "meaningful use" of electronic health record (EHR) technology and performance; starting in 2015, if they are not doing so, they will receive decreases in Medicare payments.

In addition, The Health Information Technology for Economic and Clinical Health (HITECH) Act was passed as part of ARRA, which, in part, addressed the privacy and security issues associated with the electronic transmission of health information. It also reinforced and upped the penalties possible for violations of private health information. As published in the Federal Register, "The HITECH Act was incorporated into ARRA to promote the adoption and meaningful use of health information technology."[1]

In short, the federal government provided financial incentives and penalties to physicians and other healthcare providers to start utilizing some form of health IT to electronically manage healthcare data.

Chapter 3:

The Laboratory Environment

Aggregating batches of incoming lab paper reports creates an administrative burden...

As of June 2011, there are 225,746 CLIA-certified clinical diagnostic laboratories in the United States. They range from large international corporations like Quest Diagnostics and LabCorp, to small physician office laboratories (POL) that provide a limited number of laboratory tests. The Centers for Medicare and Medicaid (CMS) and the Centers for Disease Control and Prevention (CDC) estimate that more than 7 billion laboratory tests are performed in the U.S. each year.[2] In terms of workflow, a physician will order a series of tests on a patient. That physician may be located in a private office, a clinic, or a hospital. Depending on the facility, and the type and variety of tests, the patient's sample – blood, urine, stool, or other – will go to wherever the test needs to be conducted: a reference laboratory, a hospital laboratory, a specialty laboratory, etc. Not all of these various lab service providers communicate medical data in the same format. Many still provide data on paper.

A 2011 report by the Department of Health & Human Services stated:

"Conventionally, information transactions between clinicians and laboratories (both test ordering and test result retrieval) occur via fax, mail, and dedicated phone lines and printers set up in physician offices. In some cases, patients may even act as carriers of information. There are a number of inconveniences associated with these mechanisms."

Those inconveniences include the lag time for the transfer of test results. Aggregating batches of incoming lab paper reports creates an administrative burden, and, the report cites, "paper results cannot be

easily integrated into a computerized provider record and so this mode often requires manual entry of data which is subject to human error."[3]

A single adverse drug event (ADE) in a hospital added approximately $8,750 to the cost of the hospital stay.

A 2006 report published by the Institute of Medicine of the National Academies ("Preventing Medication Errors") indicated that each year there were at least 1.5 million preventable adverse drug events (ADE) that occurred in the U.S. each year. A single ADE in a hospital added approximately $8,750 to the cost of the hospital stay. Even a low estimate of 400,000 such events annually would incur $3.5 billion of medical expenses in the hospital environment alone related to medication errors. These and other similar errors can be reduced through meaningful use of electronic healthcare data systems.

Yet the challenge, given current pressures to incorporate electronic medical records and electronic health records into all healthcare institutions, is how to deal with unstructured data and structured data that come in as external documents or outside a specific institution's electronic record format.

Chapter 4:

Data Approaches

Approximately 21 months into the ambulatory EHR installation process, about 44% of the lab results to the 2,300 faculty physicians move across an electronic interface.

Laboratories increasingly operate a laboratory information system (LIS) that provides test results directly to their clients – physicians – via a data interface. For instance, the North Shore-Long Island Jewish Health System (NSLIJHS), a major regional health system in the New York City area, operates 15 hospital sites and services a physician network that spans 200 practice sites and a growing outreach market for physicians not directly affiliated with the health system. James M. Crawford, MD, PhD, Senior Vice President for Laboratory Services, says, "Overall we perform in excess of 16 million tests system-wide. We have a central reference laboratory that does in excess of 6 million tests per year and serves as an internal reference laboratory for our hospital-based laboratories and serves as the primary site to which we direct our outreach business."

NSLIJHS is putting most of their efforts into implementing a system-wide EHR system. Obviously, with that large a system, there are complexities. For example, over the past three years the LIS platform for most of their hospital-based laboratories has been converted to Cerner Millenium. In the next several months the central reference laboratory will be upgraded from Cerner Classic to Cerner Millenium. Overlapping with these LIS upgrades are hospital activations of computerized provider order entry (CPOE) systems, which must be synchronized with the LIS platform migrations. Now, approximately 21 months into the ambulatory EHR installation process, about 44% of the lab results to the 2,300 faculty physicians move across an electronic interface. The rest are ordered and resulted through a paper fax/hand-delivery system. Crawford says he looks forward to the day when they are over 90% electronic.

So if North Shore-Long Island Jewish Health System has chosen a combination of building interfaces, a middleware solution, and a combination of EHRs and LISs, what are the other solutions?

Raymond Aller outlines four ways in which healthcare institutions tend to handle healthcare data capture. First, scan it. "They take a picture of the data and very often that's all they do with it, stick it in the archives. It's completely non-computable at that point."

If the original document is clean and made on laser-printed pages, and the scan is reasonably high-resolution, some studies indicate that OCR will "read" the words correctly greater than 98% of the time.

Thomas P. Caruso, PhD, MBA, PMP, CEO of Quantal Semantics, Inc., and an HIT consultant in Vienna, VA, agrees. "You can't do searches on that information. It's very difficult to do any research on that information. It's very difficult to figure out where the errors are or how to improve the information."

Secondly, once the data is scanned, an Optical Character Recognition (OCR) engine is applied against the scanned document in order to allow a word search. This does not necessarily involve integrating the data into an electronic health record. It allows the actual words in a scanned document to be searched, although it has limitations. For instance, the accuracy of the OCR application is dependent upon the quality of the original document and the resolution of the scan. If the original document is clean and made on laser-printed pages, and the scan is reasonably high-resolution, some studies indicate that OCR will "read" the words correctly greater than 98% of the time.

However, if the documents are faxes, multi-generational photocopies, or used an early-generation dot matrix printer, then OCR capture accuracy drops significantly and may be as low as 60 to 80%. Some graphical materials, such as lines and boxes on the original document, can confuse OCR engines.[4] Although this data is now searchable, it is not structured. It is merely indexed so it can be accessed later.

Thirdly, says Aller, "apply structure against it. For instance, someone says, this data is always on the third line of the report." Although this may sound like a reasonable way to solve the integration issue, it's not because a one-size-fits-all approach fails to recognize that laboratories and healthcare institutions do not use standardized forms where the data is "always located on the third line of the report." This results in the need for customized solutions and at least some level of manual verification.

...a one-size-fits-all approach fails to recognize that laboratories and healthcare institutions do not use standardized forms...

The fourth, of course, is that someone basically takes the original document and manually enters the data into the medical record. This is a costly and time-consuming process that can also incorporate operator entry errors.

In an ideal world, every laboratory would operate an LIS and it would be fully compatible with every institution and physician's EMR and there would be seamless communication within the systems. And that is, to some extent, the goal of current HIT regulations and incentives.

Interfaces come with their own problems. It's not just a matter of loading software or adding a telephone line. A large clinical diagnostic laboratory may have hundreds of interfaces. Laboratories may offer "custom order sets," which are batteries of tests, which complicate interfaces. A 2011 report by the Department of Health & Human Services stated:

It's clear that there is a significant gap in getting all medical data into the electronic health record...

> "Larger laboratories have initiated corporate interface verification procedures to ensure that new interfaces adhere to the company standards and regulatory requirements. To document compliance with the Clinical Laboratory Improvement Act (CLIA), labs are required to verify the interfaces of EHR vendors. The interface verification process is largely driven by interface implementation and can take as little as a month or as long as a few years, depending on the vendor and provider requirements."[3]

The report goes on to evaluate the many difficulties facing the industry in terms of interface development. "We found little evidence that the health care sector has developed clearly defined best practices around interface development. Each lab respondent described a similar approach to interface development, but it was clear that the process remains relatively loose, with large variations in the time required to set up an interface (between six weeks and six months), the level of effort required, and the exact steps that need to be deployed." They also pointed out that the costs of establishing laboratory interfaces are substantial.

When placing this into a context of over a quarter million laboratories performing over 7 billion laboratory tests annually, and adding in the cost and time issues involved with developing LIS interfaces, it's clear that there is a significant gap in getting all medical data into the electronic health record and will continue to be for some time. There is a solution.

Chapter 5:

Data Gap Solutions – Extract Systems

Extract Systems offers LabDE, a solution for automatic data capture of structured laboratory results from non-interfaced laboratories.

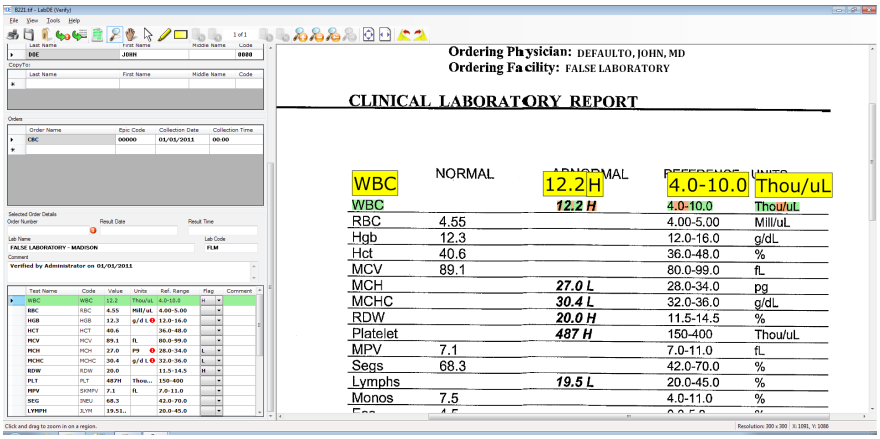
Manual healthcare data entry and merely attaching scanned documents into a medical record are not satisfactory solutions to integrating paper documents, external documents, and non-formatted structured data into the electronic medical record. That leaves scanning, the application of an OCR engine, and applying some sort of tagging or structure to that data as the most viable options. A number of companies and researchers such as MAVRO Imaging, Profdoc (Compugroup), and KnowleSys have developed approaches to this problem[5], although they often focus on a particular type of data or platform, i.e., websites or email.

Extract Systems offers LabDE, a solution for automatic data capture of structured laboratory results from non-interfaced laboratories. Anthony P. Arciniega, Business Development Healthcare at Extract Systems, says, "LabDE focuses on lab efficiency and solves real problems associated with non-interfaced external lab results. LabDE detects non-structured information, captures it, and then parses and integrates the data into our customers' information management system so the providers have all lab results at their fingertips. LabDE can send the data directly to the LIS or EMR. Customarily providers would have to search for the outpatient record in a folder somewhere as a linked image or wait for a copy to reach the patient chart. With LabDE, all outpatient results can live in their information system just like their inpatient results."

The procedure is fairly straightforward. Once the documents are scanned, the images undergo OCR. Extract Systems has developed a comprehensive and production-tested rules engine that is applied to

the document to chart lab results directly into the LIS or EMR. The company provides a custom-designed structured data solution for each customer's specific needs and unique workflows.

Extract Systems has developed a comprehensive and production-tested rules engine that is applied to the document to chart lab results directly into the LIS or EMR.



Arciniega points out that external providers will continue to be in the picture, even for health systems with extensive lab interfaces, unless there's a comprehensive HIE fix, or a way to implement an interface that's less costly and resource-intensive. "LabDE is deployed at a highly integrated health system with 400 providers and 48 clinics. They were challenged by limited IT resources and the cost required to develop an interface with every low-volume external lab. LabDE was brought in as a vendor-neutral solution to electronically capture lab results from all non-interfaced labs. The lab reports vary widely in format yet LabDE accurately detects and extracts test name, component, flag, range, units, value, collection date and time with an 88% capture rate. The data that's captured is 100% accurate." Arciniega emphasizes that there is no conflict between the use of LIS-EHR interfaces and implementing LabDE for non-interfaced labs.

In addition to applying LabDE to lab orders and pathology reports, Extract's data capture can be utilized for a number of applications involving unstructured medical data. Extract Systems also offers redaction software to protect PHI including social security/account numbers and credit card information.

Chapter 6:

Data Gap Filled

The federal government and most healthcare providers and institutions recognize the value of having medical information stored and transmitted electronically...

Although a number of healthcare institutions have attempted to become "paperless," it's clear that it will be a long time before that becomes a reality. And perhaps it never will, simply because paper is convenient and easy to work with. A 2010 survey of physician attitudes about EMRs in BMC Health Services Research[6] indicated that 58.1% of their physician respondents were skeptical that EMRs improved the quality of their medical practices. A survey released in December 2010 by the Centers for Disease Control and Prevention found that although physician use of EMRs was increasing, only 50.7% of respondents had adopted EMRs. Only 78% of patients whose physicians actually use an EMR indicated they liked the EMRs.[7] The federal government and most healthcare providers and institutions recognize the value of having medical information stored and transmitted electronically, both for cost-savings and to reduce medical errors. Patients and healthcare providers, however, apparently are not completely convinced. Also, a number of studies indicate that states, charged with developing health information exchanges (HIE), are progressing slowly, citing problems of consensus, cooperation, and financial and operational resources. Data security is also a significant challenge.[8]

Although many healthcare stakeholders have invested in EMRs, EHRs, LISs and other health IT approaches, there is still a lot of paper being generated. With 100% adoption of electronic medical records still a way off, there is clearly a need for some way to fill in the medical data gap. One effective way is to utilize an approach like Extract Systems' LabDE, which applies OCR technology and

proprietary algorithms to external documents and paper records that will allow the structured medical data to be integrated into the EHR.

Robert Aller says, "In the long term, a lot of the value of an EMR is the ability to compute the data. The least expensive method, and the one that requires the least initiative, is just leaving the information on paper somewhere and having someone try to find the paper chart. This isn't terribly useful – searching for records is a huge hidden cost. Once you put data into the computer system of any sort, it achieves availability in multiple areas, which is very important and very useful."

"In the long term, a lot of the value of an EMR is the ability to compute the data."

References

1. HITECH Act Enforcement Interim Final Rule. U.S. Department of Health & Human Services. http://www.hhs.gov/ocr/privacy/hipaa/administrative/enforcementrule/hitechenforcementifr.html
2. CLIA. Centers for Medicare & Medicaid Services. https://www.cms.gov/clia/
3. Dullabh P and Mouiduddin A. Electronic Exchange of Clinical Laboratory Information: Issues and Opportunities. http://aspe.hhs.gov/sp/reports/2009/informationissues/report.pdf
4. Optical Character Recognition (OCR): What is OCR, how does it work, and what should you expect from it? LegalScans: Paper to Searchable Databases. http://www.legalscans.com/ocr.html
5. Kristianson KJ. Data extraction from a semi-structured electronic medical record system for outpatients: A model to facilitate the access and use of data for quality control and research. Health Informatics Journal. December 2009; 15(4):305-319.
6. Boonstra A and Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Services Research 2010; 10:231.
7. Katherine Rourke. Survey Says…Patients Like EMRs (Or Think They Do). EMR and EHR. March 9, 2011. http://www.emrandehr.com/2011/03/09/survey-says-patients-want-emrs-or-think-they-do/
8. Luke Gale. Report: Obstacle-ridden course ahead for statewide HIEs. HealthImaging. February 9, 2012. http://www.healthimaging.com/index.php?option=com_articles&article=31832

Appendices

A-1
About Mark Terry

Mark Terry is a freelance writer and editor specializing in clinical diagnostics, telemedicine and biotechnology. Mr. Terry has a Bachelor of Science degree in microbiology and public health from Michigan State University. Immediately after graduating he took on the position of research assistant in an infectious disease research laboratory, then worked for 18 years in clinical genetics prior to turning to writing. Terry has published over 600 magazine and trade journal articles, 15 books and more than two dozen book-length market research reports and white papers related to clinical diagnostics. He is a member of the Association of Health Care Journalists and the Association of Genetic Technologists. Visit his website at www.markterrywriter.com.

A-2
About Extract Systems

Extract Systems is a leading provider of advanced data capture and redaction solutions for healthcare, government and other commercial sectors. Its products automatically extract and incorporate business critical content into designated information systems as structured data. This drives operational efficiency, streamlines information-rich workflows and optimizes the intelligent collection of data. Extract's award-winning redaction technology automatically detects and removes sensitive data for compliance with state and federal information security laws, including HIPAA, FOIA, Privacy and Sarbanes Oxley Act. A Microsoft Certified ISV Partner, Extract's core technology is extremely flexible and has been seamlessly integrated with health information systems, proprietary government recording systems and numerous document/content management platforms.

A-3
About DARK Daily

"Dark Daily is a concise e-news/management briefing on timely topics in clinical laboratory and anatomic pathology group management. It is a solution to the dilemma facing anyone in the laboratory profession.

DARK Daily is a concise e-news/management briefing on timely topics in clinical laboratory and anatomic pathology group management. It is a solution to the dilemma facing anyone in the laboratory profession. New developments, new technology, and changing healthcare trends make it imperative to stay informed to be successful. At the same time, the Internet, cell phones, blackberries, laptop computers and wireless devices are overwhelming any one individual's ability to absorb this crushing Tsunami of data.

DARK Daily is a quick-to-read, easy-to-understand alert on some key developments in laboratory medicine and laboratory management. It has no counterpart in the lab world. Why? Because it is produced and written by the experts at THE DARK REPORT and The Dark Intelligence Group, who know your world, understand your needs and provide you with concise, processed intelligence on only those topics that are most important to you!

You will find DARK Daily to also be an exceptionally valuable resource in laboratory and pathology management. Some of the lab industry's keenest minds and most effective experts will be offering their knowledge, their insights and their recommendations on winning strategies and management methods. Many of these experts are unknown to most lab directors. As has proven true with THE DARK REPORT for more than a decade, DARK Daily will be your invaluable— and unmatched—resource, giving you access to the knowledge and experience of these accomplished lab industry professionals.

A-4
About The Dark Intelligence Group, Inc. and THE DARK REPORT

"Membership is highly-prized by the lab industry's leaders and early adopters. It allows them to share innovations and new knowledge in a confidential, non-competitive manner.

The Dark Intelligence Group, Inc., is a unique intelligence service, dedicated to providing high-level business, management and market trend analysis to laboratory CEOs, COOs, CFOs, pathologists and senior-level lab industry executives. Membership is highly-prized by the lab industry's leaders and early adopters. It allows them to share innovations and new knowledge in a confidential, non-competitive manner. This gives them first access to new knowledge, along with the expertise they can tap to keep their laboratory or pathology organization at the razor's edge of top performance.

It offers qualified lab executives, pathologists and industry vendors a rich store of knowledge, expertise and resources that are unavailable elsewhere. Since its founding in 1996, The Dark Intelligence Group and THE DARK REPORT have played instrumental roles in supporting the success of some of the nation's best-performing, most profitable laboratory organizations.

The Dark Intelligence Group (TDIG) is headquartered in Austin, Texas. This location makes it very accessible for any laboratory organization seeking input, insight and support developing their business operations, creating effective business strategies and crafting effective sales and marketing programs that consistently generate new volumes of specimens and increasing new profits. The Dark Intelligence Group, Inc. owns and operates two Web sites in the TDIG Website network:

THE DARK REPORT

http://www.DarkReport.com

DARK Daily
Clinical Laboratory and Pathology News/Trends

http://www.DarkDaily.com

A-5
About the *Executive War College on Laboratory and Pathology Management*

Every spring since 1996, the lab industry's best and brightest gather at the *Executive War College on Laboratory and Pathology Management* to learn, to share and to network. Many consider it to be the premier source of innovation and excellence in laboratory and pathology management.

Each year, a carefully selected line-up of laboratory leaders and innovators tell the story of how their laboratories are solving problems, tackling the toughest challenges in lab medicine and seizing opportunities to improve clinical care and boost financial performance. The *Executive War College* is the place to get practical advice and solutions for the toughest lab management challenges. A unique case study format brings participants face-to-face with their most successful peers. They tell, first hand, how their laboratory solved intractable problems and successfully used new technology.

Many lab management secrets are shared, along with specific "what-not-to-do's" gained from hard-won experience! It's not pie-in-the-sky theory, but useful knowledge that can be put to use in any lab. The *Executive War College* offers superlative networking, with lab administrators and pathologists attending from countries as far away as the United Kingdom, Germany, Brazil and Australia. It makes the *Executive War College* a melting pot for all the best ideas, new lab technologies and management strategies now reshaping the laboratory industry. It's also become a recruiting ground used by headhunters and major lab organizations.

In the United Kingdom, The Dark Intelligence Group and the Association of Clinical Biochemists (ACB) have co-produced a meeting every February since 2003. Known as *Frontiers in Laboratory Medicine* (FiLM), it attracts laboratory leaders and innovators in the United Kingdom. Also featuring a case study format, this meeting pioneered the international laboratory side-by-side case study, where a North American laboratory and a United Kingdom laboratory prepare a comparison of best practices and an operational assessment of their two organizations.

In September 2005, a laboratory management meeting called *Executive Edge* was conducted in Toronto, Ontario, Canada, by The Dark Intelligence Group and QSE Consulting. It provided pathologists and lab directors in Canada with a customized meeting devoted to the strategic and operational issues of laboratory management in Canada.

A-6
About David Rasmussen

David Rasmussen is President of Extract Systems. With more than 15 years' experience leading software companies, David is currently chairman of the Board of Directors for Accelerate Madison and serves on the Greater Madison Chamber of Commerce Business Advisory Council. David earned a BA from the University of Wisconsin-Whitewater and an MBA from DePaul University in Chicago.

Terms of Use:

Contact Information

The Dark Intelligence Group, Inc.
Customer Service
800-560-6363
21806 Briarcliff Drive
Spicewood, Texas 78669

Legal Notices