Big Data ~ Big Challenge ~ Big Impact

—Understanding the Five Hurdles to Harnessing Big Data and How it Will Change the Future of Healthcare for Patients, Payers, Researchers and Governments Forever—

William A. Burns　Tomohiro Hotta

OVERVIEW:

Healthcare is in the throes of a quiet revolution, one where the steady growth of computing power and applications technology is changing the very way we create, deliver and administrator care. This confluence of factors is now making Big Data a powerful force in its own right, one that must be harnessed and understood if we are to advance. Computing in today's healthcare ecosystems has become ubiquitous, creating countless new digital puddles, lakes, tributaries, and oceans of information. A menagerie of digital devices has proliferated and gone mobile – cell phones, smart phones, laptops, and medical sensors – which in turn are generating a daily flood of new information. More healthcare business and government agencies are discovering the strategic uses of Big Data. As all these systems begin to interconnect with each other and as powerful new software tools and techniques are invented to analyze the data for valuable inferences, a radically new kind of "knowledge infrastructure" is materializing. A new era of Big Data in healthcare is emerging, and the implications for business, government, and society are enormous. This article describes today's healthcare Big Data challenges and potential impacts, and how understanding and mastering the five basic tenets of healthcare Big Data will change our environment forever.

KEY PROBLEM OF DATA FRAGMENTATION

The separation of data among labs, hospital systems, and even clinical components, like financial information and electronic health records (EHRs), serves as the main issue with leveraging healthcare data. All of these are separate repositories for information. Their sole use is to provide clinical care, or to provide scheduling information or operational information, and this fragmentation in itself is a problem if we want systems to talk to each other. Organizations can also end up with redundant information due to a maze of legacy systems. If we want to improve quality-of-care and lower cost, we need a shift from best practices to a culture of best practices – if we have them available – but also best experiences and using data from various components of health information technology (IT) to improve care and lower costs in a holistic way.

BIG DATA'S EMPHASIS ON REALTIME OR NEAR-REALTIME INFORMATION

Traditional analytics use extract, transform, and load (ETL) processes that upload data nightly or weekly to a data warehouse. Processing takes place in the warehouse, yet the trend in Big Data, at least in healthcare, is moving toward realtime or near-realtime processing. In other words, deriving value from data more immediately without waiting for batch processes. The name of the game in healthcare, in clinical decision support as well as at the point of care, is being able to understand data to make a decision. With traditional analytics, reporting focused on the past, but Big Data is more about prediction and looking forward to what may happen in the future.

PROCESSING MOVING TO WHERE DATA RESIDES

A continuing trend in healthcare information and analytics is having the processing come to the data,

instead of the other way around. Traditionally, it has been data that is moved out of a production database to a warehouse, and pulled from different repositories. At the rate data is increasing in healthcare, whether it be from medical imaging, EHRs, or other sources, moving this data around is becoming more of a challenge. So a trend is emerging whereby processing resources and system intelligence are moving to the data. This is still a very large job that is split into a number of parts and across healthcare systems that span tens if not hundreds of miles. The infrastructure knows where the data resides, and processing happens as close to the data as possible to improve performance.

SHIFT FROM "SCALE-UP" TO "SCALE-OUT"

Traditionally, information systems deployment and information processing theory have centered around a "scale-up" mentality. Yesterday's answer to these problems was always "Get me a bigger server, a more powerful server." Today, the trend is toward "scale-out." What this means is, instead of leaving older hardware nodes behind or getting rid of them, the performance and scalability of a system are improved over time by adding additional nodes. The same notion is true of information storage. Being able to scale more easily within the existing architecture (being able to add another node to the solution when needed) makes systems easier to manage. It is solutions like these that the industry is moving toward as an alternative to the "rip and replace" mentality.

DEMAND FOR SaaS FROM SMALLER ORGANIZATIONS

Among today's healthcare organizations, the trend that is perhaps gaining the most momentum is smaller hospitals and research organizations leaning toward software-as-a-service (SaaS). While EHR vendors generally install the solution on behalf of clients and customers, they also provide applications, specifically analytics, to add complementary additional value. A lot of these analytics solutions continue to focus on meaningful use and quality metrics. It is anticipated that these trends will accelerate over time toward SaaS, especially for smaller

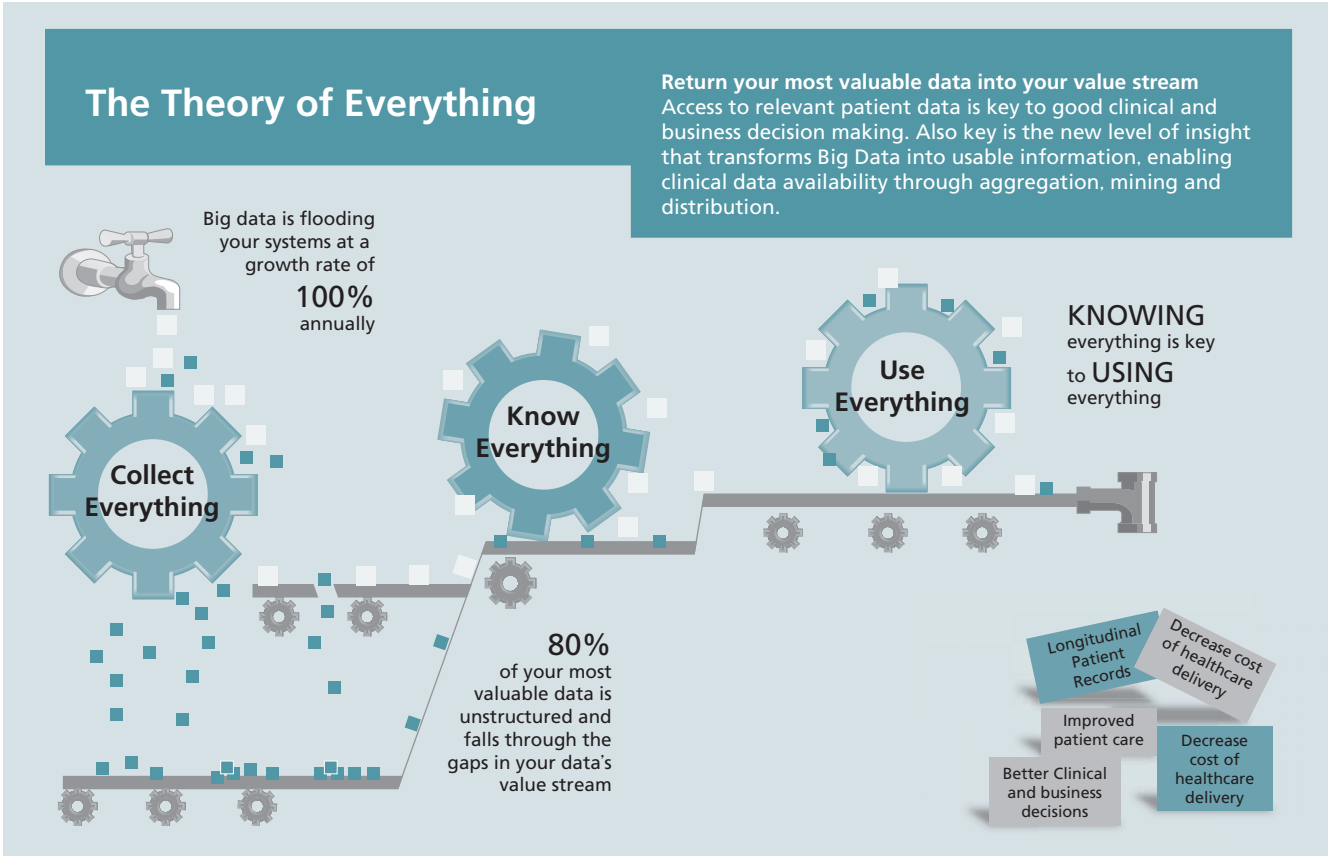

Fig. 1 | Transforming Enterprise Information
Enterprise and market knowledge are encapsulated across everything in today's modern enterprise. Unlocking this value can create tremendous competitive advantages for companies.

organizations, as it makes sense for them to look at hosted analytics solutions and hosted services.

HITACHI CLINICAL REPOSITORY (HCR)

For hospitals with a vendor-neutral archive (VNA) strategy, Hitachi Clinical Repository (HCR) acts as a VNA (See **Fig. 2**). It supports import and storage of Digital Imaging and Communications in Medicine (DICOM)*1 studies as well as query and retrieval of studies, regardless of the original picture archiving and communication system (PACS) vendor. HCR, a registered medical device, provides a single online repository that enables protection, search, and retrieval across all image and non-image healthcare data, and provides scalable images, information, and lifecycle management in an imaging environment. HCR enables images and related information to be queried, stored, and retrieved in a manner defined by open standards while maintaining patient privacy and security. HCR provides a patient-centric approach that transcends upgrades and changes to viewing, acquisition, and workflow-management components. Such components are interchangeable, without the need to migrate, convert, or change the data formats or interfaces.

Support for Open Standards

HCR fully supports DICOM and Health Level Seven (HL7)*2, the most common accepted medical protocols. Support for these standards means other systems in an enterprise can query HCR to retrieve data. In addition, metadata is captured in extensible markup language (XML), the ideal format for easy retrieval by the various applications that may be used to access and view the data, such as EHRs or physician portal applications.

Assured Patient Privacy and Security

Security is a critical requirement of any healthcare IT infrastructure. HCR provides multiple layers of features designed to safeguard and secure the data entrusted to it. It is a highly scalable, secure, self-healing, and cloud-enabled object repository platform that is capable of supporting multiple simultaneous applications. HCR

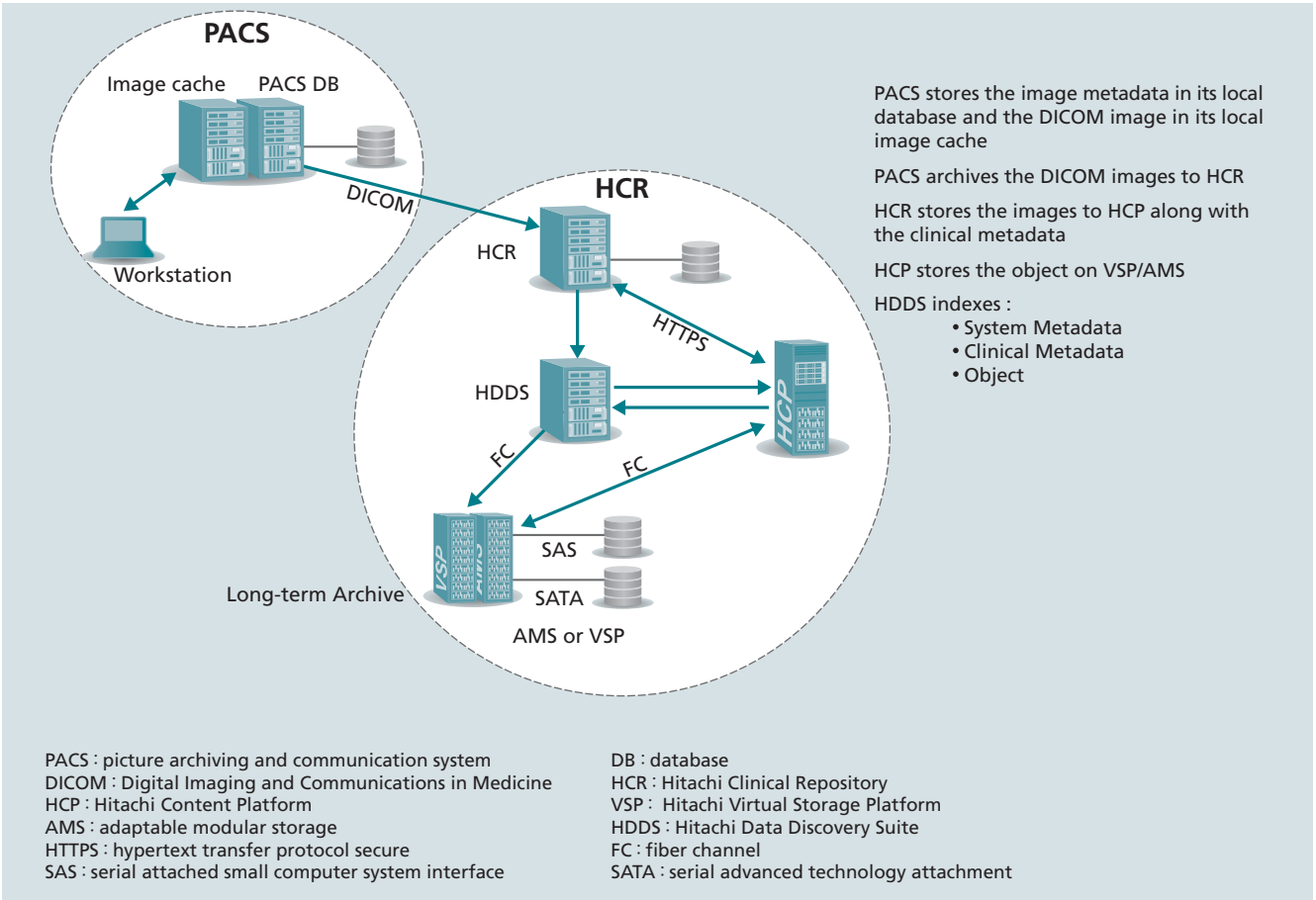

Fig. 2 | Creating Enterprise Medical Information
Solutions like HCR enable organizations to separate departmental data from source applications, creating a new world of enterprise data assets.

takes a layered approach to security, ensuring the safety of data while restricting unauthorized access. Stored data is encrypted at rest to prevent tampering, and HCR provides automated policy enforcement, failover, and integrity checks to ensure that data remains secure and robust.

Intelligent Data Lifecycle Management

Managing the flow of data and information, from creation and initial storage to obsolescence, is the goal of a data lifecycle management strategy. It is also a requirement of good regulatory standards. HCR intelligent data lifecycle management automates and optimizes data storage, taking into account clinical usage patterns. HCR manages the tiering of data between different disk performance levels with robust retention and deletion capabilities.

Maintaining Data Interoperability without DICOM Wrappers

When converting unstructured data (which is typically non-DICOM) to structured data, a VNA typically relies on DICOM wrapping. This involves creating DICOM headers that contain the metadata elements to include with the data object. HCR captures data in its native format, indexes the metadata associated with the study, and returns a uniform resource locator (URL) to tell the consuming applications where the medical information is stored. Maintaining the original format of the data allows it to be recalled by the original application and modified without conversion, while maintaining data interoperability.

CONCLUSIONS

Big Data presents many exciting opportunities to improve today's modern healthcare systems. There are incalculable opportunities to make scientific research more productive, and to accelerate discovery and innovation around new healthcare delivery paradigms. People can use new tools to help improve their health and well-being, and medical care can be made more efficient and effective. Government, too, has a great stake in leveraging Big Data to improve the delivery of government health services and to monitor for threats to national security (such as syndromic surveillance and bioterrorism). New Big Data technologies are also likely to arise to help people make sense of an otherwise bewildering flood of information.

However, Big Data also presents many formidable challenges to governments and citizens precisely because data technologies are becoming so pervasive, intrusive, and difficult to understand. How shall society protect itself against those who would misuse or abuse our healthcare information? What new regulatory systems, private-law innovations, or social practices will be capable of controlling these behaviors and how should we even define what is socially and legally acceptable when the practices enabled by Big Data are so novel and often arcane? These are some of the important open questions posed by the rise of Big Data in our healthcare systems. One way or another, society will need to take some innovative, imaginative leaps to ensure that our healthcare information technologies and techniques are used effectively and responsibly.

---

*1 DICOM is the registered trademark of the National Electrical Manufacturers Association for its standards publications relating to digital communications of medical information.

*2 Health Level Seven and HL7 are registered trademarks of Health Level Seven International.

REFERENCES

(1) R. Bryant et al., "Big-Data Computing: Creating Revolutionary Breakthroughs in Commerce, Science and Society," http://www.cra.org/ccc/docs/init/Big_Data.pdf (Dec. 2008).

(2) C. Anderson, "The End of Theory: The Data Deluge Makes the Scientific Method Obsolete," Wired Magazine,16.07,http://www.wired.com/science/discoveries/magazine/16-07/pb_theory (Jun. 2008).

(3) J. Manyika et al., "Big data: The next frontier for innovation, competition, and productivity," McKinsey Global Institute (May 2011).

(4) "Materials Genome Initiative for Global Competitiveness," National Science and Technology Council (Jun. 2011).

(5) R. Platt et al., "The New Sentinel Network—Improving the Evidence of Medical-Product Safety," New England Journal of Medicine, pp. 645-647 (Aug. 2009).

ABOUT THE AUTHOR

William A. Burns
Vice President,
Global Health & Life Sciences
Hitachi Data Systems

Joined Hitachi in 2008 and leads the Global Health & Life Sciences Team at Hitachi Data Systems. Mr. Burns has extensive experience in the digital healthcare arena, including point-of-care disease management systems, diagnostic imaging, ambulatory patient monitoring, clinical research platforms, and regulatory compliance. He has proven instrumental in setting the leadership foundation in the development of technology enhanced clinical and business initiatives for Hitachi Data Systems clients.

Tomohiro Hotta
Strategy Planning & Development Office
Hitachi, Ltd.

Joined Hitachi, Ltd. in 2000, and now works at the Platform Strategy, Strategy Planning & Development Office. He is currently engaged in the development of healthcare business.