Making Sense of your Data

The Data Center Vision:
Making Sense of the Data

David L. Brock

The Data Center, Massachusetts Institute of Technology, Building 35, Room 234, Cambridge, MA 02139-4307, USA

ABSTRACT

The Data Center at the Massachusetts Institute of Technology is a new initiative charged with researching and developing the languages, protocols and technologies to seamless integrate data and models across the global network. The Center will develop the infrastructure, recommend the standards and build the prototype applications that enable the interoperation of data and analytic models within and across enterprises. The technologies and standards created by the Center will be open and freely distributed. This paper presents the vision and goals of the Data Center – and lays the foundation for a new *Intelligent Information Network*.

About the Author

David L. Brock is Principal Research Scientist at the Massachusetts Institute of Technology, and co-founder and a Director at the Auto-ID Center (now EPCGlobal, Inc. and Auto-ID Laboratories). The Center was an international research consortium formed as a partnership among more than 100 global companies and five leading research universities. David is also Assistant Research Professor of Surgery at Tufts University Medical School and Founder and Chief Technology Officer of endoVia Medical, Inc., a manufacturer of computer controlled medical devices. Dr. Brock holds bachelors' degrees in theoretical mathematics and mechanical engineering, as well as master and Ph.D. Degrees, from MIT. David can be reached at dlb@mit.edu

1.0 INTRODUCTION

The Data Center at the Massachusetts Institute of Technology is a new initiative charged with researching and developing the languages, protocols and technologies to seamless integrate data and models across the global network. The intent is to truly "make sense of all the data" which is now so readily available. Specifically, the Center will develop the infrastructure, recommend the standards and build prototype applications that enable the interoperation of data and analytic models within and across enterprises. The technologies and standards created by the Center will be open and freely distributed. This paper presents the vision and goals of the Data Center – and lays the foundation for a new *Intelligent Information Network*.

2. PROBLEM

We live in a time of unprecedented information. The global internet has truly transformed the world, so that thoughts and ideas – in fact the wealth of human knowledge – are now shared instantly across the globe. In addition, governments, organizations and companies routinely manage equal amounts of information internally through their secure networks.

This is only the beginning. Newer technologies such as radio frequency identification (RFID) and sensor networks are poised to flood the network with even greater amounts of information. Corporations are now struggling to effectively manage and integrate their legacy data systems with this new real-time information.

The difficultly is not in the quantity and availability of information, but in the interpretation of this information to make intelligent decisions. This lies at the heart of the Data Center's mission – to convert seemingly meaningless data into information for action.

3. VISION

The Data Center envisions a world in which in information is shared and automatically harmonized among disparate data sources – and further this information is combined with mathematical models to produce meaningful data for intelligent decisions. More precisely, we intend multiple, disparate databases – or portions thereof – to merge effortlessly *without* any a-priori agreement on content. Furthermore these combined data would communicate with analytic models that provide analyzes, estimates, predictions and plans. Finally, we envision models *themselves* would combine with other models to form larger synthetic systems.

The realization of this vision would have vast benefit to all organizations that deal with information – which is *everyone*. Although broad in scope, this vision can be achieved – albeit incrementally – by using today's technology to provide a foundation and creating new infrastructure to build tomorrow's *intelligent information network*.

4. STRATEGY

Our strategy for achieving this vision is composed of four parts. First, we will use existing and emerging standards where possible to address near-term issues in data synchronization and sensor integration. Second, we will research and develop new languages and protocols to address fundamental problems in data and model interoperation.   The "M" Language is in the early stages of being such a solution. Third, we will explore new technologies and methods to manage, analyze and visualize information and algorithm. Grid computing, immersive visualization and multi-modal user interfaces are a few of the techniques we will examine. Finally, we will test and evaluate all of these advances against real-world business concerns and practical use-cases. Taken together these elements present a multi-faceted approach to the deeper issues of information management and interoperable analytic modeling.

5. CONCLUSION

This vision of freely interchangeable models and data is a dream we believe can be achieved – and one that stands to revolutionize the way we manage information. The Data Center is dedicated to researching and developing the languages, protocols and technologies needed to achieve this goal and to deliver practical solutions to academia and industry.