inxight

Building Successful Enterprise Business Intelligence Solutions

Ramana Rao   HP World 2002
CTO & Senior Vice President   Sept 25, 2002
Inxight Software, Inc.

http://www.ramanarao.com

Getting Value from Content

This presentation focuses on innovative solutions that harness the vast resources of unstructured data. The audience will learn:

1. Why unstructured data presents challenges for business intelligence

2. How to enhance and increase employee productivity by improving users' information discovery experience

Abstract

Each year, corporations worldwide spend over $30 billion purchasing electronically published proprietary information. U.S. companies alone spend at least $100 billion on internally authored documents (e.g., product literature, research papers, etc.). Unfortunately, this high-value content, so-called unstructured data, ends up not being used. This presentation focuses on innovative solutions that harness the vast resources of unstructured data so that corporations can produce precise, relevant, and timely information to drive their mission-critical decisions and business processes.

Ramana Rao is senior vice president, chief technology officer and a founder of Inxight Software. Ramana leads the charge on Inxight's product vision and architecture and is particularly passionate about providing real solutions to customer problems. Ramana serves as the company's evangelist and speaks at numerous conferences annually and is frequently interviewed and profiled by media and analysts on design, innovation, knowledge management, and information access technologies.

Throughout his career, Ramana has designed and evangelized software that aims to extend the intellectual and creative reach of knowledge workers. At Xerox's Palo Alto Research Center [PARC] for ten years, Ramana performed pioneering work in intelligent information access, digital libraries, information visualization, and user interfaces. His work includes 25 patent filings and numerous refereed research papers. He received bachelor's and master's degrees in computer science and engineering from the Massachusetts Institute of Technology.

Preliminary Talk Outline (final slides available at http://www.ramanarao.com**)**

1) Introduction

2) Problem: Content is underutilized

    a) Enterprises buy & create lots of content. Meanwhile, people can't effectively find, analyze or utilize this information. Thus valuable content is underutilized in business decision-making and planning

3) Enterprise Realities & Requirements

    a) Classic Search and Browsing are commonly deployed, but each have well-understood limitations. For example, Search requires you to know what you are looking for, while browsing pages is a very ineffective way to navigate large information spaces.

    b) Portal and content mgmt. solutions don't address content organization or access

    c) Manual organization and tagging of content is expensive and doesn't scale very well. Completely automated solutions can't work.

    d) Internet Standards-based Integration into complex enterprise infrastructure is essential to minimize IT costs and risks.

4) Review of Specific Content Access Technologies

    a) Classic and Enhanced Search and Browse – Search and browsing can be integrated in many ways. One powerful way is to show the results of search directly with visualizations.

    b) Categorization and Taxonomy – Organizing documents into hierarchical classification systems enable more efficient searching and browsing.

    c) Information Extraction – Extracting entities, concepts, facts and relationship from text allows for automatically tagging content with metadata

    d) Information Visualization – Visual representations of content collection can increase the bandwidth of interaction so users can understand the structure of collections and navigate to where they want to more quickly.

5) Demonstrations

    a) Discovering and managing taxonomies

    b) Enhanced Search & Browse end user experience

6) Conclusions

7) Additional resources at http://www.ramanarao.com The purpose of this site is to share articles, talks, and thoughts on *Information Flow*, intelligent approaches to information access that make people, organizations, and society more productive and more creative.

    a) *Information Flow monthly email newsletter*, browse archive of past issues and subscribe at www.ramanarao.com

    b) *The Secret of Unstructured Data*, e-doc magazine, Sept/Oct 2002, available from www.edocmagazine.com

    c) *See and Go Manifesto*, Interactions magazine, Aug 1999, paper on next generation user interface