

Intelligent-based Data Mining for Space Sciences

H. Karimabadi

SciberQuest, Inc.

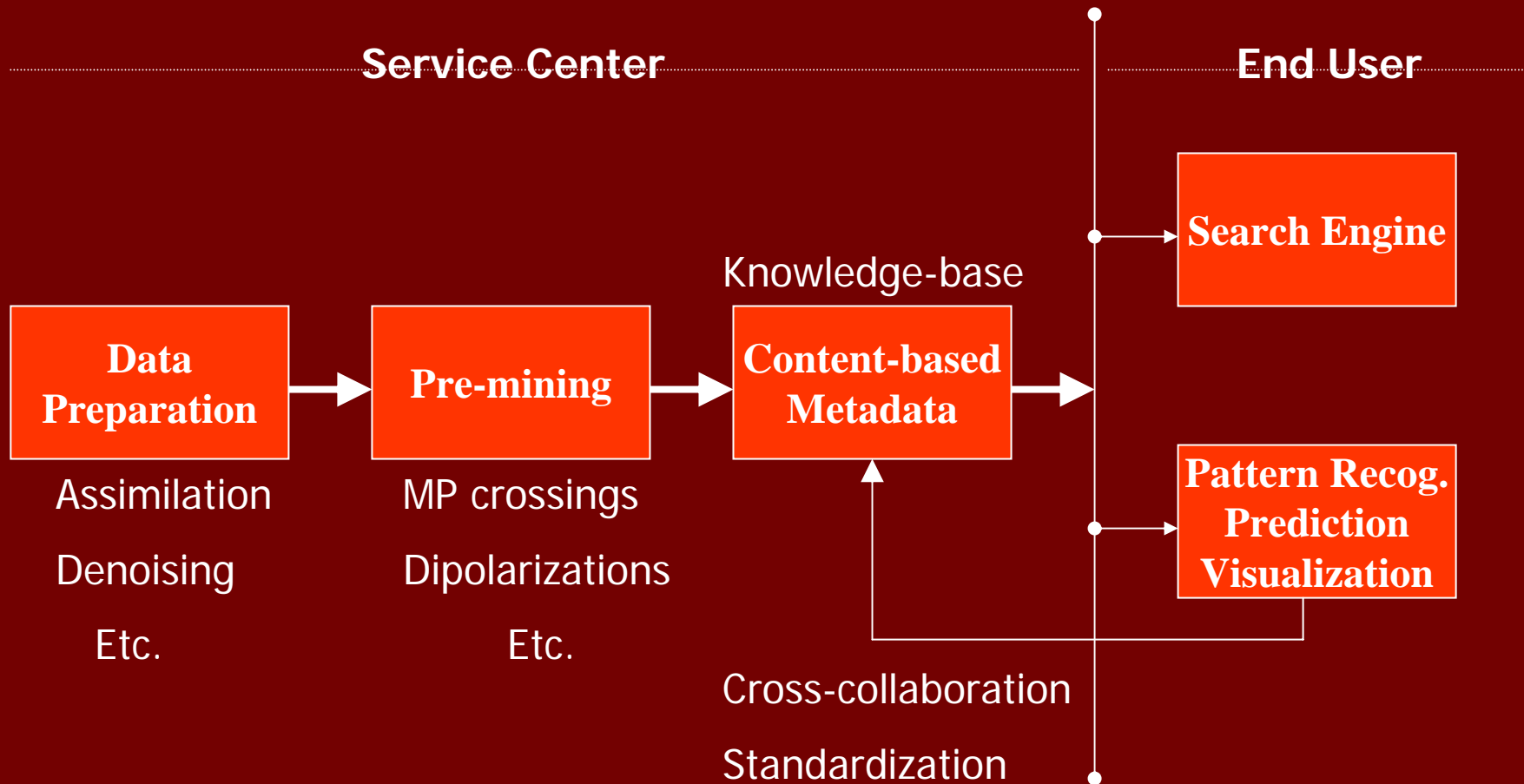
A New Perspective on How Systems Evolve: Discrete Systems

- Traditional means of modeling physical systems is through a continuous time-stepped methodology
- A discrete methodology based on events is proving more advantageous
- Event driven methodology offers a new way to decipher spatio and temporal relations in the data

What is Data Mining?

- Data/image preparation (e.g., image denoising, intelligent filtering, edge detection, etc.)
- Content-based metadata, and data storage
- Pattern / event recognition
- Classification
- Event prediction
- Variable dependency discovery
- Event screener / matcher
- Intelligent data driven search engine
- Visualization of statistical inferences

Schematic of Our Proposed Infrastructure



Data Mining Applications

- Mining the Spacecraft Data
- Mining the Simulation Data
- Embedded intelligent agents in Simulations (e.g., as predictor of mesh refinement in AMR)

Data Mining Techniques

- Bayesian techniques
- Neural networks
- Evolutionary algorithms
- Support Vector Machines
- Relevant Input Processor Network (RIPNet)

Unique Challenges of Data Mining in Space Sciences

- Dispersed and non-standardized data formats and
- Simultaneous analysis of data from multiple spacecraft
- System too complex for off-the-shelf data analysis packages
- Even commercial packages require expertise for proper use

Our Approach – Phase I

Develop a customized software by:

- Use simulation data for training
- Use single and multi-spacecraft data
- Work directly with experimentalists for GUI design, and development of basic modules

Features of Our Solution

- Desktop as well as client-server delivery – connectivity to data portals
- Parallel performance for large-scale statistical studies, multi-user access, pre-mining of NASA datasets, etc.
- “Groupware” capability to allow for collaboration as well as standardization of basic modules (e.g., defining what constitutes an FTE, etc.)
- Content-based metadata

Research and Development Concepts – Phase II

- Use of discrete event methodology to derive dependencies and evolutionary rules
- Image/data preparation based on discrete technology
- Intelligent agents:
 - Self regulating algorithms (interacts with users to ensure correct use of algorithm, provide answers)
 - Autonomous operation (self-mining)