
Sapphire System Architecture

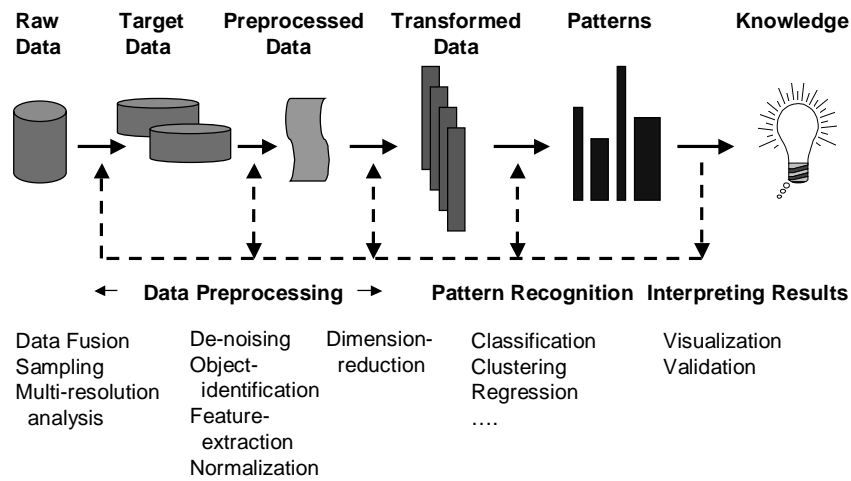
Chandrika Kamath
Center for Applied Scientific Computing
Lawrence Livermore National Laboratory
<http://www.llnl.gov/casc/people/kamath>
January 16, 2002

UCRL-PRES- 146654 : This work was performed under the auspices of the U.S. Department of Energy by the University of California Lawrence Livermore National Laboratory under contract No. W-7405-Eng-48.

Sapphire: a research project in scientific data mining

- **The focus of Sapphire is three-fold**
 - **research in robust, accurate, scalable algorithms**
 - **incorporating the research into parallel, portable, software modules within a flexible system architecture**
 - **application of the software to practical problems**
- **Analyze data from**
 - **simulations**
 - **experiments**
 - **observations**
- **Details, including publications, at**
 - **<http://www.llnl.gov/casc/sapphire>**

The Sapphire view of the end-to-end data mining process



CK/3

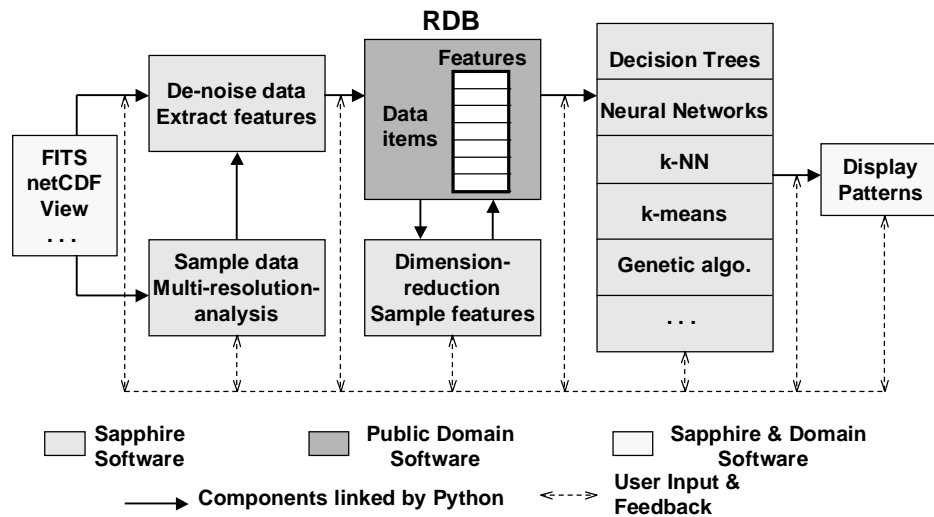
Our software is designed to address the diverse needs of our applications

- Not all problems require the entire process
- Not all algorithms are suitable for a problem
- Algorithms typically depend on several parameters
- Intermediate data must be handled appropriately
- Domain dependent and independent parts must be clearly identified
- Should be able to accommodate a growing data set

→ Our system design enables us to conduct research and develop software in the context of several problems.

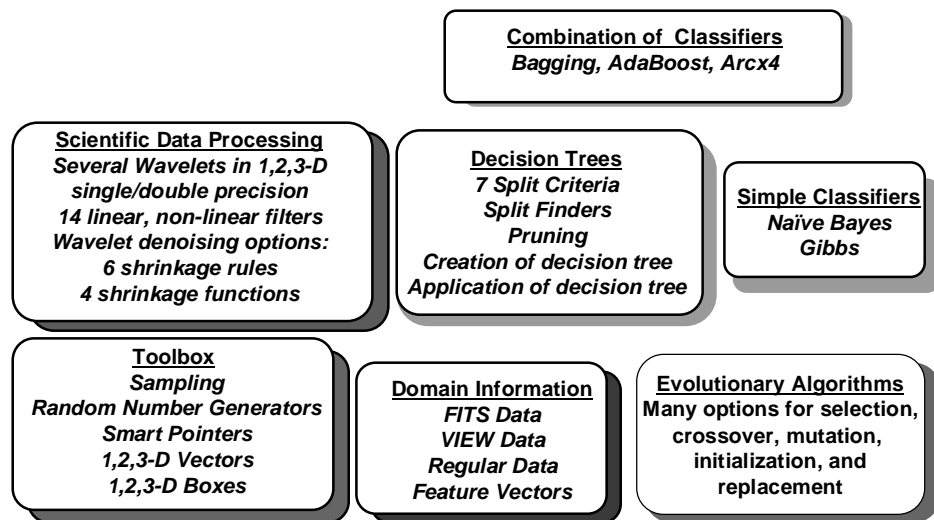
CK/4

The Sapphire approach: a flexible, portable, scalable system architecture



CK/5

Sapphire software: Version 1.0.0 released in September 2001 (C++, serial version)



CK/6

“Non-mathematical” lessons learned

- **Understanding the data and formulating the problem takes time**
- **Must not under-estimate time to**
 - **obtain access to the data**
 - **read, write, and display the data**
 - **bring the data into a consistent format**
- **Obtaining consistently labeled data is non-trivial**
- **Handling a growing data set can be tricky**