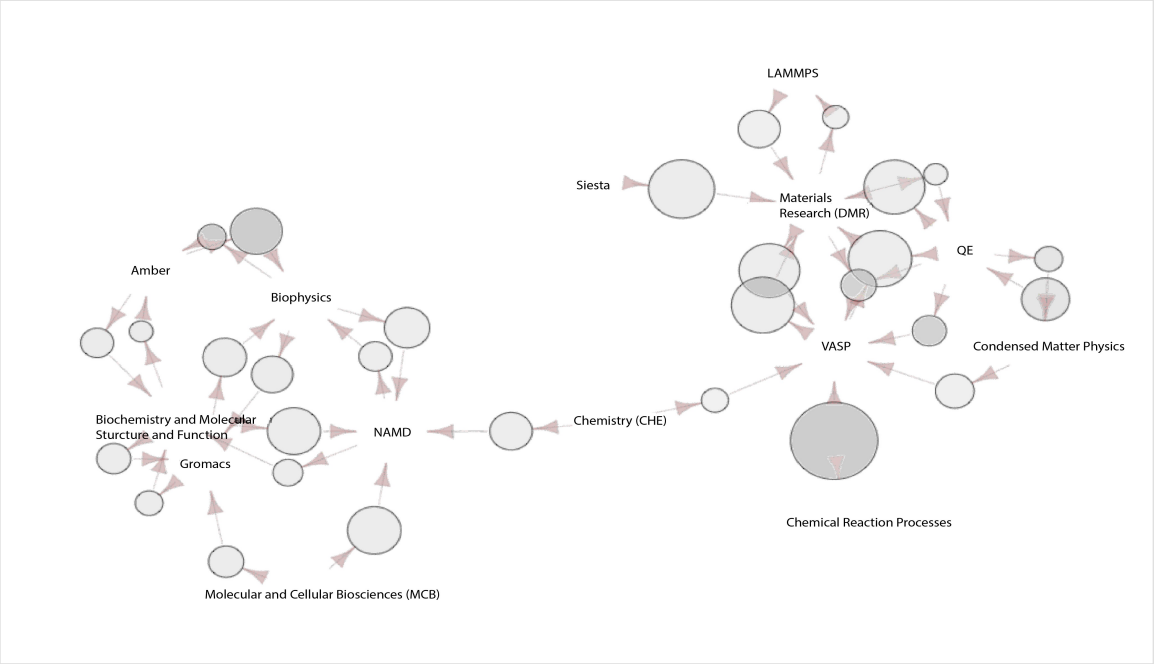
XALT Dataset: Proof of Concept Graphics
Images and statistical analysis by Dr. Weijia Xu.



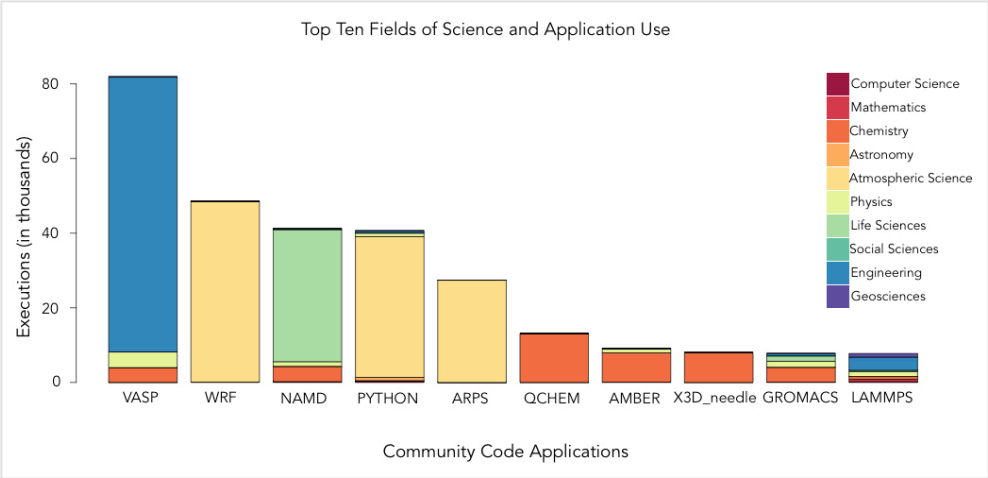
The graphic above, produced from data from 2014/07 to 2015/06, is the product of statistical analysis to identify job patterns. Using associative analysis, Dr. Xu used the data to visualize the relations of scientific fields through the use of community codes.

Association analysis can identify inference rules that can predict the occurrence of one set of variables based on the co-occurrence of another set of variables. The method consists of two steps: step one identifies objects that commonly occur together as a frequent item set; and step two identifies inference rules among objects based on their presence in the frequent item set.

The graphic above shows the most relevant patterns found in the data, the fields of science and applications are connected via directional links. The direction of the link indicates the inference from one to the other, and in most cases there are links in both directions between pairs. The size of the circle in the middle of each link indicates the strength of the inference, the bigger the circle, the stronger the inference.

The most revealing pattern shows that users from the field of Chemistry use applications that are also commonly used by Biology and Physics users. The connection is not causal and is established based on executable usage by fields of science. It shows direct connections between field of science and executable, and potential indirect connections among executables. The results indicate that Chemistry users use both VASP (for Quantum Mechanics, Chemistry), and NAMD (Nanoscale Molecular Dynamics for Biophysics), with certain level of significant frequency, and that VASP and NAMD are also significantly used by users from related fields of science. Such connections are at the tip of understanding the interdisciplinary nature of the computational research conducted on Stampede.

XALT Dataset: Proof of Concept Graphics
Images and statistical analysis by Dr. Weijia Xu.



The bar graph shows the statistical distribution of self-selected fields of science over the top ten community applications used on Stampede, expressed as number of executions in one year. The graph shows that applications that they considered domain specific such as VASP (Vienna Ab initio Simulation Package for Molecular Dynamics in Chemistry), Gromacs (for Biomolecular Dynamics), and LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator for soft and hard materials Molecular Dynamics), are actually being shared across various scientific disciplines.